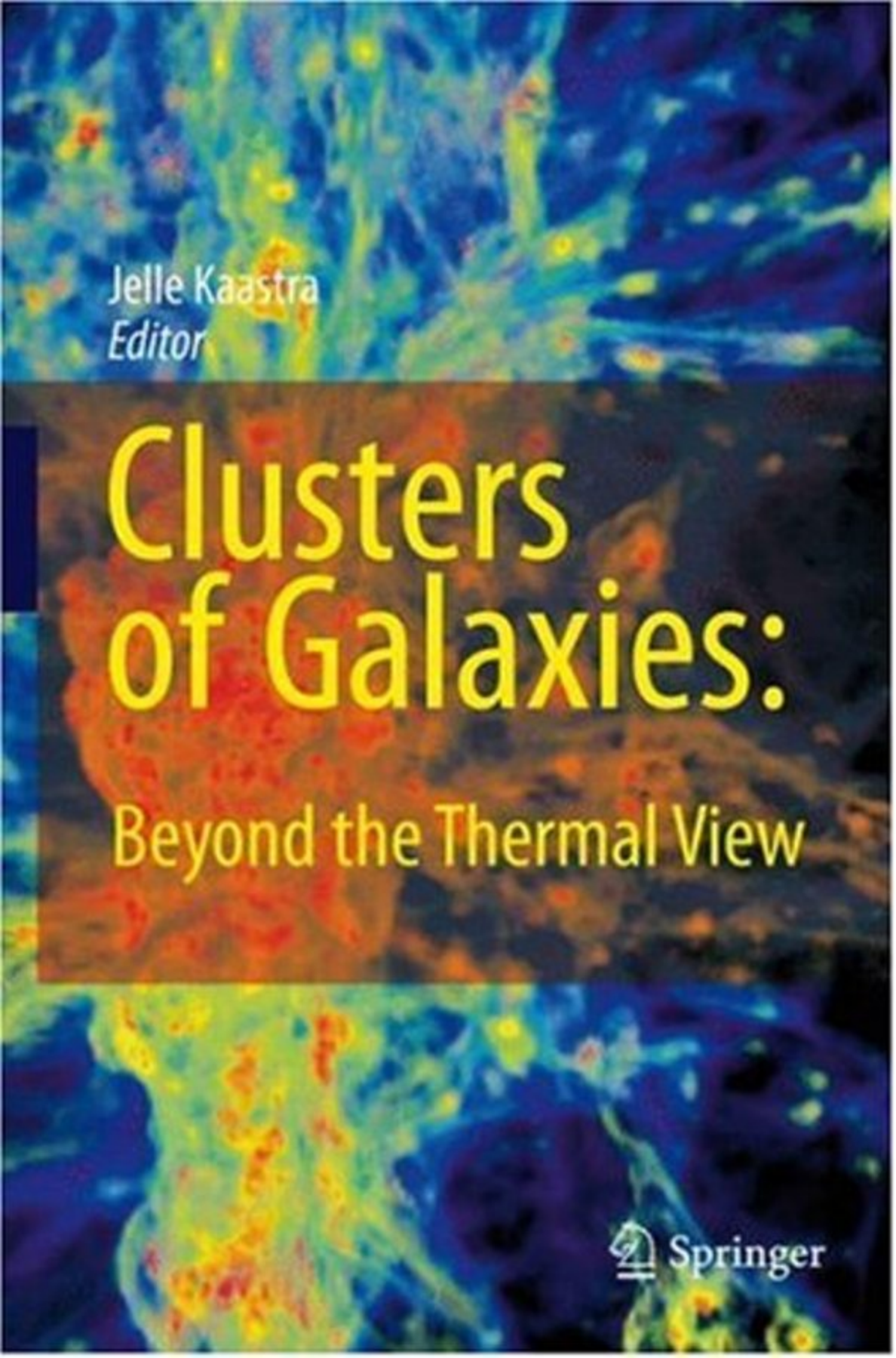
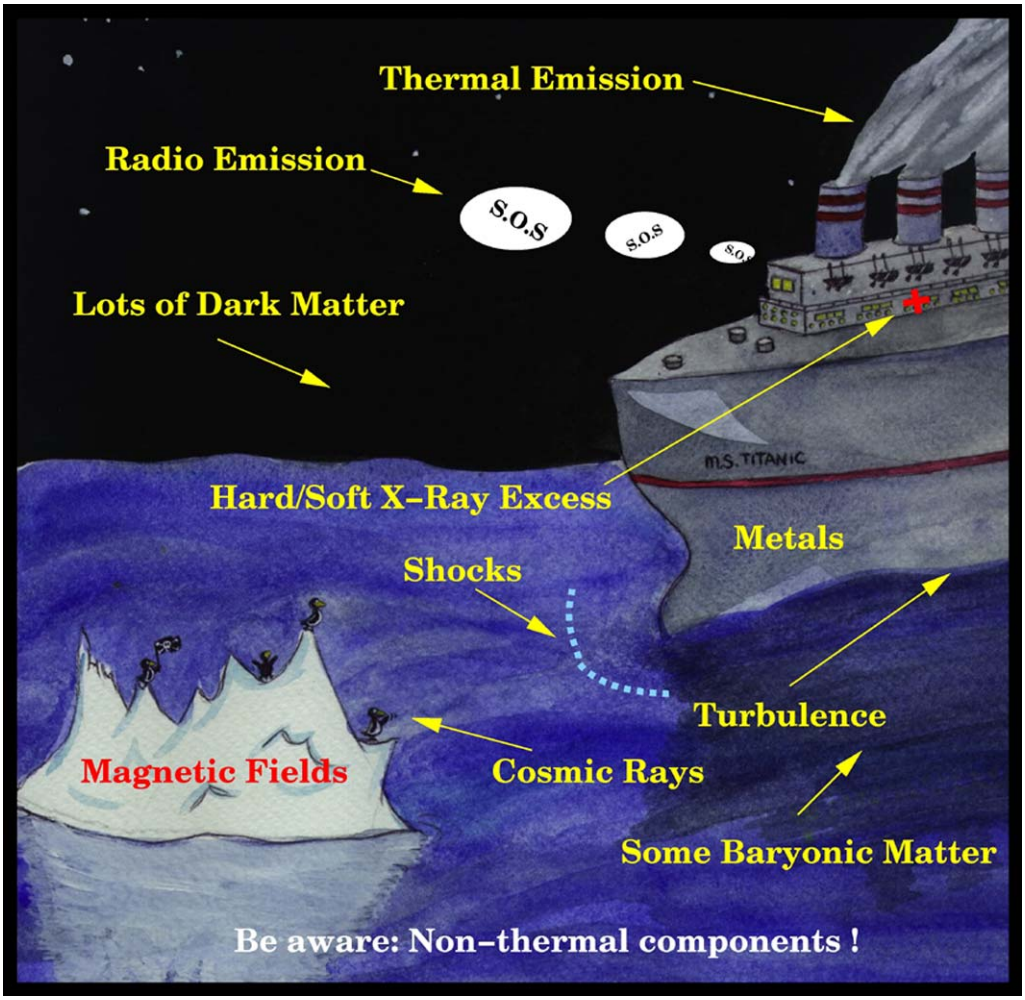

Jelle Kaastra
Editor
Clusters of Galaxies:
Beyond the Thermal View
Springer


Thermal Emission
Radio Emission
S.O.S
S.O.S
S.O.S
Lots of Dark Matter
M.S. TITANIC
Hard/Soft X-Ray Excess
Metals
Shocks
Turbulence
Magnetic Fields
Cosmic Rays
Some Baryonic Matter
Be aware: Non-thermal components !

Jelle Kaastra
Editor

Clusters of Galaxies: Beyond the Thermal View

Previously published in *Space Science Reviews* Volume 134, Issues 1–4, 2008

Springer

Jelle Kaastra
Space Research Organisation Netherlands (SRON)
Utrecht, The Netherlands

Cover illustration: Background image kindly provided by F. Vazza, 2008.

Frontispiece: Courtesy of Klaus Dolag, Max Planck Institut für Astrophysik, Garching, Germany.

Library of Congress Control Number: 2008925092

ISBN-978-0-387-78874-6 e-ISBN-978-0-387-78875-3

Printed on acid-free paper.

1

springer.com

Contents

Chapter 1
Clusters of Galaxies: Beyond the Thermal View

J.S. Kaastra · A.M. Bykov · S. Schindler · J.A.M. Bleeker · S. Borgani · A. Diaferio · K. Dolag · F. Durret · J. Nevalainen · T. Ohashi · F.B.S. Paerels · V. Petrosian · Y. Rephaeli · P. Richter · J. Schaye · N. Werner

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9326-3

Abstract We present the work of an international team at the International Space Science Institute (ISSI) in Bern that worked together to review the current observational and theo-

J.S. Kaastra (✉) · J.A.M. Bleeker · F.B.S. Paerels · N. Werner
SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands
e-mail: j.kaastra@sron.nl

J.A.M. Bleeker
e-mail: j.a.m.bleeker@sron.nl

N. Werner
e-mail: n.werner@sron.nl

J.S. Kaastra · J.A.M. Bleeker
Astronomical Institute, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, The Netherlands

A.M. Bykov
A.F. Ioffe Institute of Physics and Technology, St. Petersburg, 194021, Russia
e-mail: byk@astro.ioffe.ru

S. Schindler
Institute for Astro- and Particle Physics, University of Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria
e-mail: Sabine.Schindler@uibk.ac.at

S. Borgani
Department of Astronomy, University of Trieste, via Tiepolo 11, 34143 Trieste, Italy

S. Borgani
INAF—National Institute for Astrophysics, Trieste, Italy

S. Borgani
INFN—National Institute for Nuclear Physics, Sezione di Trieste, Trieste, Italy
e-mail: borgani@oats.inaf.it

A. Diaferio
Dipartimento di Fisica Generale "Amedeo Avogadro", Università degli Studi di Torino, via P. Giuria 1, 10125, Torino, Italy

retical status of the non-virialised X-ray emission components in clusters of galaxies. The subject is important for the study of large-scale hierarchical structure formation and to shed light on the "missing baryon" problem. The topics of the team work include thermal emission and absorption from the warm-hot intergalactic medium, non-thermal X-ray emission in clusters of galaxies, physical processes and chemical enrichment of this medium and clusters of galaxies, and the relationship between all these processes. One of the main goals of the team is to write and discuss a series of review papers on this subject. These reviews

A. Diaferio
Istituto Nazionale di Fisica Nucleare (INFN) Sezione di Torino, via P. Giuria 1, 10125 Torino, Italy
e-mail: diaferio@ph.unito.it

K. Dolag
Max-Planck-Institut für Astrophysik, P.O. Box 1317, 85741 Garching, Germany
e-mail: kdolag@mpa-garching.mpg.de

F. Durret
Institut d'Astrophysique de Paris, CNRS, UMR 7095, Université Pierre et Marie Curie,
98bis Bd Arago, 75014 Paris, France
e-mail: durret@iap.fr

J. Nevalainen
Observatory, P.O. Box 14, 00014 University of Helsinki, Helsinki, Finland
e-mail: jnevalai@astro.helsinki.fi

T. Ohashi
Department of Physics, School of Science, Tokyo Metropolitan University, 1-1 Minami-Osawa,
Hachioji, Tokyo 192-0397, Japan
e-mail: ohashi@phys.metro-u.ac.jp

F.B.S. Paerels
Department of Astronomy and Columbia Astrophysics Laboratory, Columbia University,
550 West 120th Street, New York, NY 10027, USA
e-mail: frits@astro.columbia.edu

V. Petrosian
Department of Applied Physics, Stanford University, Stanford, CA 94305, USA

V. Petrosian
Kavli Institute of Particle Astrophysics and Cosmology, Stanford University, Stanford, CA, 94305, USA
e-mail: vahep@stanford.edu

Y. Rephaeli
School of Physics & Astronomy, Tel Aviv University, Tel Aviv 69978, Israel

Y. Rephaeli
Center for Astrophysics and Space Sciences, University of California, San Diego, La Jolla,
CA 92093-0424, USA
e-mail: yoelr@wise.tau.ac.il

P. Richter
Institut für Physik, Universität Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany
e-mail: prichter@astro.physik.uni-potsdam.de

J. Schaye
Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands
e-mail: schaye@strw.leidenuniv.nl

are intended as introductory text and reference for scientists wishing to work actively in this field. The team consists of sixteen experts in observations, theory and numerical simulations.

Keywords Galaxies: clusters: general · Intergalactic medium · Large-scale structure of universe · X-rays: galaxy clusters

1 Scientific Rationale of the Project

Clusters of galaxies are the largest gravitationally bound structures in the Universe. Their baryonic composition is dominated by hot gas that is in quasi-hydrostatic equilibrium within the dark matter dominated gravitational potential well of the cluster. The hot gas is visible through spatially extended thermal X-ray emission, and it has been studied extensively both for assessing its physical properties and also as a tracer of the large-scale structure of the Universe.

Clusters of galaxies are not isolated entities in the Universe: they are connected through a filamentary cosmic web. Theoretical predictions indicate the way this web is evolving. In the early Universe most of the gas in the web was relatively cool ($\sim 10^4$ K) and visible through numerous absorption lines, designated as the so-called Lyα forest. In the present Universe, however, about half of all the baryons are predicted to be in a warm phase (10^5–10^7 K), the Warm-Hot Intergalactic Medium (WHIM), with temperatures intermediate between the hot clusters and the cool absorbing gas causing the Lyα forest.

The X-ray spectra of clusters are dominated by the thermal emission from the hot gas, but in some cases there appears to be evidence for hard X-ray tails or soft X-ray excesses. Hard X-ray tails are difficult to detect, and one of the topics for the team is a discussion on the significance of this detection (yet contradictory) in existing and future space experiments. Various models have been proposed to produce these hard X-ray tails, and our team reviews these processes in the context of the observational constraints in clusters.

While in some cases soft excesses in clusters can be explained as the low-energy extension of the non-thermal hard X-ray components mentioned above, there is evidence that a part may also be due to thermal emission from the WHIM. The signal seen near clusters then originates in the densest and hottest parts of the WHIM filaments, where the accelerating force of the clusters is highest and heating is strongest. A strong component of this emission is line radiation from highly ionised oxygen ions, and the role of this line emission and its observational evidence will be reviewed.

WHIM filaments not only can be observed because of their continuum or line emission, but also through absorption lines if a sufficiently strong continuum background source is present. The evidence for absorption in both UV and X-ray high-resolution spectra is discussed. Future space missions will be well adapted to study these absorption lines in more detail.

In particular in absorption lines the lower density parts between clusters become observable. In these low density regions of the WHIM not only collisional ionisation but also photo-ionisation is an important process. In general, the physics of the WHIM is challenging due to its complexity since there are many uncertain factors including the heating and cooling processes, the chemical enrichment, the role of supernova-driven bubbles or starburst winds, ram-pressure stripping, the role of shocks, magnetic fields, etc. More detailed (and sophisticated) hydrodynamical simulations with state-of-the-art spatial (and temporal) resolution are required in order to follow the impact of some (if not all) of these important processes. In particular chemical enrichment is an important process to consider as it leads to many observable predictions. We review the various physical processes relevant for the WHIM, the methods that are used to simulate this and the basic results from those models.

2 Timeliness of this Work

The first detections of non-virialised components in or between clusters of galaxies such as thermal emission or absorption from the WHIM or the presence of hard X-ray tails have now been made. There are several initiatives for new space missions to study the physics of the WHIM, either in emission or absorption, from the USA, Japan, Italy and The Netherlands, and new missions for studying hard X-ray tails are being designed in France, Japan and the USA. Significant theoretical progress is being made in this field, also thanks to the enormous leap in computing power for numerical models. In addition to the considerable intrinsic interest in the astrophysics of clusters, these systems are fundamental probes of the underlying cosmology and of the large-scale structure. This is the time to put the expertise of observers and theoreticians in different fields together, in order to review our current knowledge and make it available to the community in a self-contained and comprehensive—yet concise—review volume.

3 Organisation of the Team Work

Our team consisted of sixteen members; two members unfortunately could not attend both meetings; for one participant a replacement could be found. The people involved in this international team were Xavier Barcons (Spain), Johan Bleeker (Netherlands, co-organiser), Andrei Bykov (Russia, co-organiser), Stefano Borgani (Italy), Antonaldo Diaferio (Italy), Klaus Dolag (Germany), Florence Durret (France), Jelle Kaastra (Netherlands, organiser), Jukka Nevalainen (Finland), Takaya Ohashi (Japan), Frits Paerels (USA), Vahé Petrosian (USA), Yoel Rephaeli (Israel), Philipp Richter (Germany), Joop Schaye (Netherlands), Sabine Schindler (Austria, co-organiser), and Norbert Werner (Netherlands). In addition a few people outside the team helped in writing the review papers or even have taken up the lead in writing these papers: Serena Bertone (UK), Chiara Ferrari (Austria), Federica Govoni (Italy), Dunja Fabjan (Italy), Luca Tornatore (Italy), and Rob Wiersma (Netherlands).

Our team started with a one week meeting at the end of October 2006 at ISSI in Bern. Each team member introduced a topic, and in the following discussion the relevant issues for the review were collected. After the discussions, subteams were formed around the different topics, and the outlines of the different chapters as presented in the present volume were defined. Each team member participated in several subteams and papers, either as first author or co-author. In the time between the first and the second meeting, several subteams were able to work out these drafts further. At the second meeting, the progress was reviewed, comments were made on the drafts and the remaining time was used to work further on the drafts.

After the second meeting, the draft papers were finished by the authors and sent to two internal referees from the team who are non-experts in the field, in order to check the comprehensibility of the papers and of course for the normal work of any referee. All papers were then reviewed by an anonymous, external and expert referee. That there was a lively discussion and interaction is proven by the more than 700 emails that were exchanged between the authors and the editor between July and December 2007.

4 About the Title of the Book

Our series of review papers both appears as a special volume of Space Science Reviews, as well as in the form of a book published by Springer. One of the hardest items to resolve

was to find a catchy title for the book. At our last meeting in June 2007 we had a long and lively discussion about it without a clear outcome. Only shortly before the deadline we converged to the title of the book (which was taken also as the title of this chapter/paper). While discussing what is beyond the thermal view, we spend of course also quite some time on thermal emission. It is obvious that without a thorough understanding of the thermal aspects of the cluster gas, it is not well possible to go beyond. But in this series of reviews we hope to demonstrate that cluster physics is a rich subject, and that there is more beyond galaxies, dark matter, and "just" hot gas with only one interesting parameter, its temperature (well, and its emission measure), needed to estimate cluster masses. We go beyond this, and present non-thermal emission, shocks, magnetic fields, chemical enrichment, and all the other interesting processes shown on the front cover of the book. Also, we go beyond the—somewhat diffuse—physical boundaries of clusters to the warm-hot intergalactic medium with all its interesting astrophysical aspects. We hope that our book and the individual papers help the readers also to go beyond what is written here, and we invite them to explore the topics of this book further and to advance our understanding of them.

5 Final Remarks

The team members look back to an interesting and rewarding project. The time spent at ISSI has been successful, both from a point of view of the project, the interactions and the atmosphere surrounding the project. The help and assistance of ISSI and its staff members has been invaluable for this work.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". SRON is supported financially by NWO, the Netherlands Organization for Scientific Research. A.M.B. acknowledges the RBRF grant 06-02-16844, a support from RAS Presidium Programs, and a support from NASA ATP (NNX07AG79G). S.S. acknowledges financial support by the Austrian Science Foundation (FWF) through grants P18523-N16 and P19300-N16, by the Tiroler Wissenschaftsfonds and through the UniInfrastrukturprogramm 2005/06 by the BMWF. A.D. gratefully acknowledges support from the PRIN2006 grant "Costituenti fondamentali dell'Universo" of the Italian Ministry of University and Scientific Research and from the INFN grant PD51. F.D. acknowledges support from CNES. J.N. acknowledges support from the Academy of Finland. F.P. acknowledges support from the Dutch Organization for Scientific Research NWO. N.W. acknowledges support by the Marie Curie EARA Early Stage Training visiting fellowship.

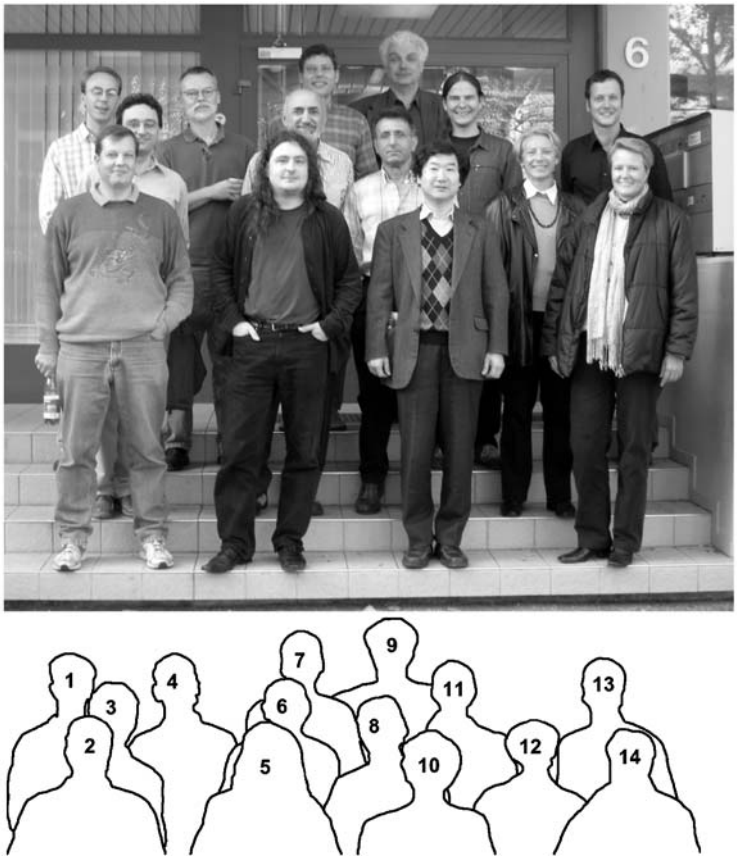


Fig. 1 The team members in front of the ISSI building at the first meeting: Joop Schaye (*1*), Klaus Dolag (*2*), Antonaldo Diaferio (*3*), Frits Paerels (*4*), Jukka Nevalainen (*5*), Vahé Petrosian (*6*), Jelle Kaastra (*7*), Yoel Rephaeli (*8*), Andrei Bykov (*9*), Takaya Ohashi (*10*), Norbert Werner (*11*), Florence Durret (*12*), Philipp Richter (*13*) and Sabine Schindler (*14*)

Chapter 2
Clusters of Galaxies: Setting the Stage

A. Diaferio · S. Schindler · K. Dolag

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9324-5

Abstract Clusters of galaxies are self-gravitating systems of mass $\sim 10^{14}$–$10^{15} h^{-1}\,\mathrm{M}_{\odot}$ and size $\sim$ 1–$3h^{-1}$ Mpc. Their mass budget consists of dark matter ($\sim 80\%$, on average), hot diffuse intracluster plasma ($\lesssim 20\%$) and a small fraction of stars, dust, and cold gas, mostly locked in galaxies. In most clusters, scaling relations between their properties, like mass, galaxy velocity dispersion, X-ray luminosity and temperature, testify that the cluster components are in approximate dynamical equilibrium within the cluster gravitational potential well. However, spatially inhomogeneous thermal and non-thermal emission of the intracluster medium (ICM), observed in some clusters in the X-ray and radio bands, and the kinematic and morphological segregation of galaxies are a signature of non-gravitational processes, ongoing cluster merging and interactions. Both the fraction of clusters with these features, and the correlation between the dynamical and morphological properties of irregular clusters and the surrounding large-scale structure increase with redshift.

In the current bottom-up scenario for the formation of cosmic structure, where tiny fluctuations of the otherwise homogeneous primordial density field are amplified by gravity, clusters are the most massive nodes of the filamentary large-scale structure of the cosmic web and form by anisotropic and episodic accretion of mass, in agreement with most of the observational evidence. In this model of the universe dominated by cold dark matter, at the present time most baryons are expected to be in a diffuse component rather than in stars and

A. Diaferio (✉)
Dipartimento di Fisica Generale "Amedeo Avogadro", Università degli Studi di Torino, Via P. Giuria 1, 10125, Torino, Italy
e-mail: diaferio@ph.unito.it

A. Diaferio
Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Torino, Via P. Giuria 1, 10125, Torino, Italy

S. Schindler
Institut für Astro- und Teilchenphysik der Leopold-Franzens Universität Innsbruck, Technikerstrasse 25/8, 6020, Innsbruck, Austria

K. Dolag
Max-Planck-Institut für Astrophysik, Karl-Schwarzschildstr. 1, P.O. Box 1317, 85741, Garching b. München, Germany

galaxies; moreover, $\sim 50\%$ of this diffuse component has temperature ~ 0.01–1 keV and permeates the filamentary distribution of the dark matter. The temperature of this Warm-Hot Intergalactic Medium (WHIM) increases with the local density and its search in the outer regions of clusters and lower density regions has been the quest of much recent observational effort.

Over the last thirty years, an impressive coherent picture of the formation and evolution of cosmic structures has emerged from the intense interplay between observations, theory and numerical experiments. Future efforts will continue to test whether this picture keeps being valid, needs corrections or suffers dramatic failures in its predictive power.

Keywords Galaxies: clusters: general · Intergalactic medium · Cosmology: large-scale structure of Universe

1 Preamble

The present chapter provides the general framework of the physics of clusters of galaxies: it outlines how clusters connect to several astrophysical issues, from cosmology to the formation of galaxies and stars. Some of these topics are discussed in the chapters of this volume; we refer to these chapters when appropriate. When the topic we mention here is not dealt with elsewhere in this volume, we refer to some of the most recent reviews.

2 The Observational Framework

2.1 Cluster Components

2.1.1 Mass Content

The discovery of clusters as concentrations of galaxies on the sky dates back to 1784, when Charles Messier mentions the Virgo cluster in his *Connaissance des Temps* (see the fascinating historical review by Biviano 2000). The first systematic optical surveys of clusters appeared almost two centuries later (Abell 1958; Zwicky et al. 1961–1968).

In 1930s, Zwicky, by assuming that the cluster in the Coma constellation, containing hundreds of bright galaxies (Fig. 1),[1] is in virial equilibrium, found that the mass required to bind the system gravitationally should be roughly hundred times larger than the sum of the masses of the individual galaxies (Zwicky 1933, 1937). In fact, if the cluster is an isolated, spherically symmetric system in dynamical equilibrium, the virial theorem yields the total mass

$$M = \frac{3\sigma^2 R}{G}, \tag{1}$$

where G is the gravitational constant, σ is the dispersion of the galaxy velocities along the line of sight, and R is the cluster size

$$R = \frac{\pi}{2}\frac{N(N-1)}{2}\left(\sum_i \sum_{i>j} \frac{1}{r_{ij}}\right)^{-1}, \tag{2}$$

[1] See Biviano (1998) for a historical perspective on the Coma cluster.

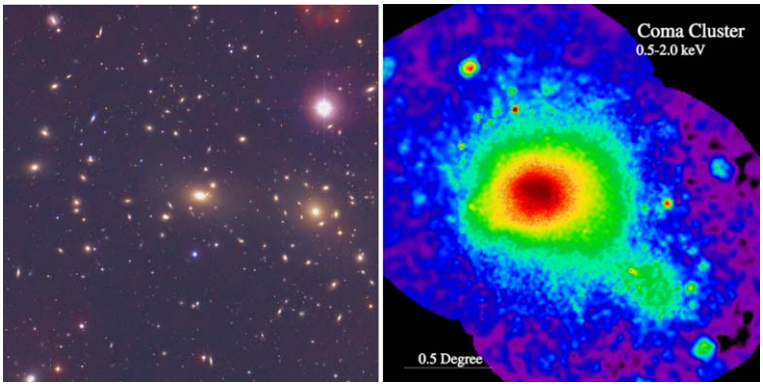


Fig. 1 *Left panel*: Optical image of the central part of the Coma cluster (23 × 23 arcmin) centered on the galaxy NGC4889; the bright galaxy on the right is NGC4874. *Right panel*: X-ray image (ROSAT) of Coma (164 × 150 arcmin) (from http://chandra.harvard.edu/photo/2002/0150/more.html)

where r_{ij} is the projected separation between galaxies i and j, and N is the number of galaxies; (2) assumes that all the galaxies have the same mass. Clusters are not isolated systems and the correct application of the virial theorem requires the inclusion of a surface term (The and White 1986; Girardi et al. 1998). However, correcting for this factor, and correcting also for departures from spherical symmetry, for galaxies of different masses and for the presence of interlopers, namely galaxies which are not cluster members but appear projected onto the cluster field of view, is not sufficient to fill in the discrepancy between the virial mass and the sum of the galaxy masses. We can reach the same conclusion with a more sophisticated dynamical analysis based on the Jeans equations (Biviano 2006). Zwicky's result was the first indication of the existence of dark matter.

With the advent of astrophysics from space in 1960s, the *Uhuru* satellite revealed that clusters are the most luminous extended X-ray sources on the sky (Gursky et al. 1972, Fig. 1). The origin of the X-ray luminosity L_X was very early interpreted as thermal Bremsstrahlung emission from a hot intracluster plasma (Felten et al. 1966; Bykov et al. 2008b—Chap. 8, and Kaastra et al. 2008—Chap. 9, this issue):

$$L_X = \int n_{\rm e}(\mathbf{r}) n_{\rm ions}(\mathbf{r}) \Lambda[T(\mathbf{r})] {\rm d}^3\mathbf{r}, \tag{3}$$

where $n_{\rm e}$ and $n_{\rm ions}$ are the electron and ion number densities in the intracluster medium (ICM) and $\Lambda(T)$ is a cooling function. For temperatures $kT \gtrsim 2$ keV (k is the Boltzmann constant), when the ICM is almost fully ionised, we have $\Lambda(T) \propto T^{1/2}$. The existence of a hot diffuse X-ray emitting gas implies the presence of a deep gravitational potential well that keeps the gas confined. This evidence made untenable the hypothesis that clusters are chance superpositions of galaxies or expanding systems (Ambartsumian 1961). By assuming hydrostatic equilibrium and spherical symmetry, the cumulative mass within radius r is

$$M(<r) = -\frac{kTr}{G\mu m_{\rm p}} \left(\frac{{\rm d}\ln \rho_{\rm gas}}{{\rm d}\ln r} + \frac{{\rm d}\ln T}{{\rm d}\ln r} \right), \tag{4}$$

where μ is the mean molecular weight, $m_{\rm p}$ the proton mass and $\rho_{\rm gas}$ the gas mass density. This type of analysis confirmed that at least 70% of the cluster mass is dark (Cowie et al.

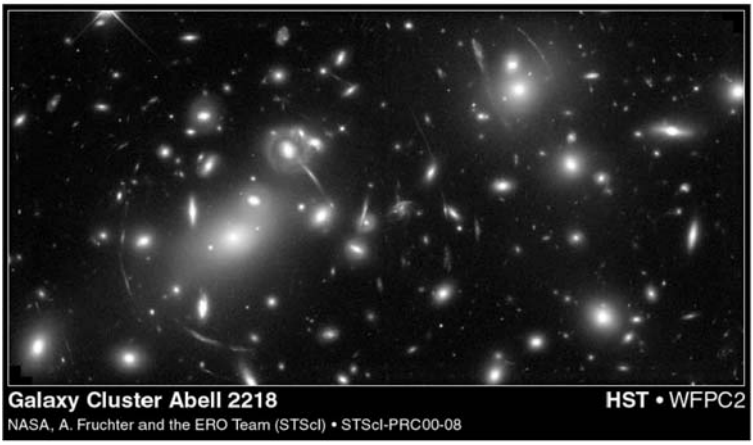


Fig. 2 HST optical image of Abell cluster A 2218. Many arcs and arclets are the images of background galaxies distorted by the gravitational field of the cluster (from http://antwrp.gsfc.nasa.gov/apod/ap011007.html)

1987), and the rest is almost all in the diffuse ICM component: galaxies only contribute a few percent to the total cluster mass, as inferred by Zwicky in 1937.

In this same 1937 paper, Zwicky also realized that one way to avoid the assumption of dynamical equilibrium when estimating the cluster mass is to use the deflection of light of background sources determined by the cluster gravitational potential well (Fig. 2); for an axially symmetric cluster, the total mass projected along the line of sight within the distance r of closest approach of the light rays to the cluster centre is

$$M(<r) = \frac{rc^2}{4G}\alpha, \tag{5}$$

where c is the speed of light, and α is the deflection angle (e.g. Schneider 2006). The first indication of a gravitational lensing effect in a galaxy cluster appeared fifty years later (Lynds and Petrosian 1986). To date, both strong gravitational lensing (Hennawi et al. 2008), which creates multiple images of background sources, and weak gravitational lensing (Dahle 2007), which produces small induced ellipticity in the shape of background galaxies, indicate that, if General Relativity is correct on the scales of clusters, ∼80–90% of the cluster mass is dark.

An alternative method to estimate the cluster mass, both in their virial region and their outskirts, which does not rely on the virial equilibrium assumption, is the caustic technique introduced by Diaferio and Geller (1997) and Diaferio (1999). This technique exploits the distribution of galaxies in redshift space to infer the galaxy escape velocity from the gravitational potential well of the cluster, and hence its mass. In fact, in the redshift diagram, namely line-of-sight velocity vs. projected distance r from the cluster centre, cluster galaxies populate a trumpet-shaped region limited by two curves, named caustics, whose separation $\mathcal{A}(r)$ at each radius r is proportional to the escape velocity. For a spherically symmetric cluster, the mass within radius r is

$$M(<r) = \frac{1}{2G}\int_0^r \mathcal{A}^2(x)\mathrm{d}x. \tag{6}$$

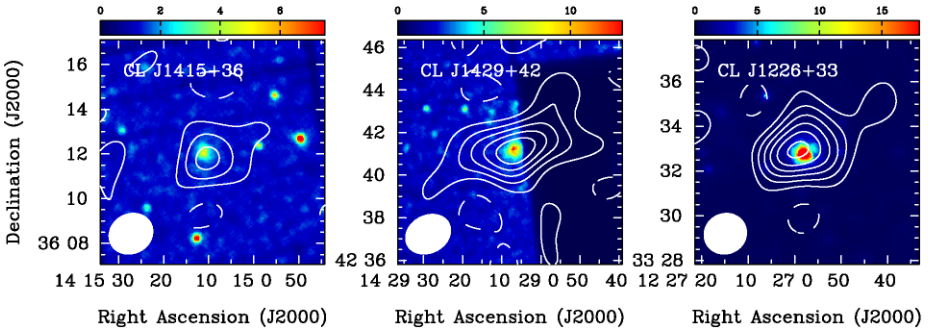


Fig. 3 Contours of the SZ effect measurements with SZA overlaid on the X-ray images (XMM EPIC/MOS) of the clusters ClJ1415.1 + 3612 ($z = 1.03$), ClJ1429.0 + 4241 ($z = 0.92$) and ClJ1226.9 + 3332 ($z = 0.89$). The *filled ellipses* show the FWHM of the SZ observations. See Muchovej et al. (2007) for details

All the mass estimation methods used to date, and described here, indicate that the dark matter contributes ∼ 80–90% of the total cluster mass, the ICM contributes ∼ 10–20%, and the galaxies contribute less than a few percent. With the advance of X-ray spectroscopy, it has become clear that the ICM is chemically enriched to ∼ 0.5 Solar abundances (Werner et al. 2008—Chap. 16, this issue): given the relative mass contribution of the cluster components, it follows that the total mass in metals in the ICM is larger than the sum of the mass in metals of the individual galaxies.

2.1.2 Clusters as Probes for Cosmology

Right after the detection of the X-ray emission from clusters, Sunyaev and Zeldovich (1972) realised that inverse Compton scattering of the Cosmic Microwave Background (CMB) photons by the free electrons of the ICM can produce secondary anisotropies in the CMB maps: $\Delta I_{\rm CMB}/I_{\rm CMB} = f(\nu)y$, where $f(\nu)$ is a spectral function and

$$y = \int \frac{kT}{m_{\rm e}c^2}\sigma_{\rm T} n_{\rm e} {\rm d}l \tag{7}$$

is the Comptonization parameter; here, $m_{\rm e}$ is the electron mass, $\sigma_{\rm T}$ the Thomson cross-section, and the integral is along the line of sight l. Typical ICM temperatures kT and electron number densities $n_{\rm e}$ in clusters yield $\Delta I_{\rm CMB}/I_{\rm CMB} \approx 10^{-4}$ (Fig. 3). Only over the last few years the sensitivity of radio telescopes and the control of systematics have been sufficient to measure reliably this Sunyaev-Zeldovich (SZ) effect and its spectral signature in massive clusters (Bonamente et al. 2006). The major advantage of the SZ effect over other cluster observables is that the amplitude of the CMB fluctuation is independent of the cluster redshift z, because the CMB intensity increases as $(1 + z)^4$ and compensates the $(1 + z)^{-4}$ decrease of the SZ "surface brightness". This property has recently boosted an entire research field of cluster surveys with the SZ effect (Carlstrom et al. 2002). In fact, "flux-limited" SZ cluster surveys are almost mass-limited cluster surveys (Bartlett 2006). This property is extremely relevant, because the mass of clusters is a sensible cosmological probe, as we explain below.

If clusters are in virial equilibrium, we can derive simple relations between their global properties, namely mass, galaxy velocity dispersion, number of galaxies (richness), X-ray luminosity, ICM temperature, and so on. By considering for example the virial relation

$3kT/(2\mu m_{\rm p}) = GM/R$, the scaling relation between the total mass M and the gas temperature kT reads

$$kT = 3.229 \left(\frac{\mu}{0.6}\right) \left(\frac{\delta}{500}\right)^{1/3} \left(\frac{M}{10^{14} h^{-1}\,{\rm M}_\odot}\right)^{2/3} {\rm keV} \tag{8}$$

where δ is the average cluster overdensity w.r.t. the critical density $3H_0^2/(8\pi G)$ of the universe, with $H_0 = 100h$ km s^{-1} Mpc^{-1} being the Hubble constant at the present time. Quantities in (8) are normalized to typical cluster values. Analogously, we can write the total X-ray luminosity

$$L_X = 1.327 \times 10^{43} \left(\frac{f_{\rm gas}}{0.1 h^{-3/2}}\right)^2 \left(\frac{0.6}{\mu}\right) \left(\frac{n}{10^{-3} h^2\,{\rm cm}^{-3}}\right) \left(\frac{T}{\rm keV}\right)^{0.4} \times \left(\frac{M}{10^{14} h^{-1}\,{\rm M}_\odot}\right) h^{-2}\,{\rm erg\,s}^{-1}, \tag{9}$$

where we approximated the cooling function $\Lambda(T) = 0.843 \times 10^{-23} (kT/{\rm keV})^{0.4}$ erg cm^3 s^{-1} at $kT \gtrsim 1$ keV, which holds for gas with poor metallicity, and assumed $n_{\rm e} = n_{\rm ions} \equiv n = f_{\rm gas}\rho/(\mu m_{\rm p})$, where $f_{\rm gas}$ is the fraction of the cluster total mass in the ICM and ρ is the cluster total mass density.

These scaling relations provide a straightforward method to estimate the cluster mass from a directly observable quantity, like the X-ray temperature. Quantities related to X-ray observations have received particular attention because of their robustness and the relatively simple physics involved in their correlations (Rosati et al. 2002).

Cluster masses are extremely relevant to constrain the cosmological model, because clusters populate the exponential tail of the mass function of galaxy systems. If the power spectrum of the primordial perturbations of the mass density field in the early universe is a power law with index n, the number per unit volume $[{\rm d}n(M)/{\rm d}M]{\rm d}M$ of galaxy systems with total mass in the range $(M, M + {\rm d}M)$ is (Press and Schechter 1974)

$$\frac{{\rm d}n(M)}{{\rm d}M}{\rm d}M = \frac{1}{\sqrt{\pi}} \frac{\bar{\rho}}{M^2} \left(1 + \frac{n}{3}\right) \left(\frac{M}{M_*}\right)^{(n+3)/6} \exp\left[-\left(\frac{M}{M_*}\right)^{(n+3)/3}\right] {\rm d}M, \tag{10}$$

where $\bar{\rho}$ is the (constant) comoving mean mass density of the universe and M_* is a parameter depending on the normalisation σ_8 of the power spectrum and on the structure growth factor, which, in turn, depends on time, the cosmological density parameter Ω_m and the cosmological constant Ω_Λ. M_* increases with time when $n > -3$ and $M_* \sim 10^{14} h^{-1}\,{\rm M}_\odot$ at the present epoch. Clearly, since the exponential cut off dominates the mass function at $M \gtrsim M_*$, the evolution of the cluster number density is a very sensitive indicator of the power spectrum normalisation and of the cosmological parameters (Fig. 4). In practice, the application of this idea requires modern versions of the Press-Schechter mass function which are more sophisticated than (10) (Sheth and Tormen 1999); moreover, care must be taken to deal with the degeneracy among the cosmological parameters (e.g. Voit 2005), especially between $\Omega_{\rm m}$ and σ_8, which follow a relation of the form $\sigma_8 = A\Omega_{\rm m}^{-\alpha}$, with $A \approx 0.5$ and $\alpha \approx 0.5$ (e.g. Reiprich and Böhringer 2002). Nevertheless, by simply assuming that the mean mass density of the universe must be smaller than the mean density within clusters and larger than the mean density computed including only the richest clusters, Abell already in 1965 estimated $\Omega_{\rm m} \approx 0.2$, a value remarkably close to the currently accepted estimate $\Omega_{\rm m} = 0.26$ (Spergel et al. 2007).

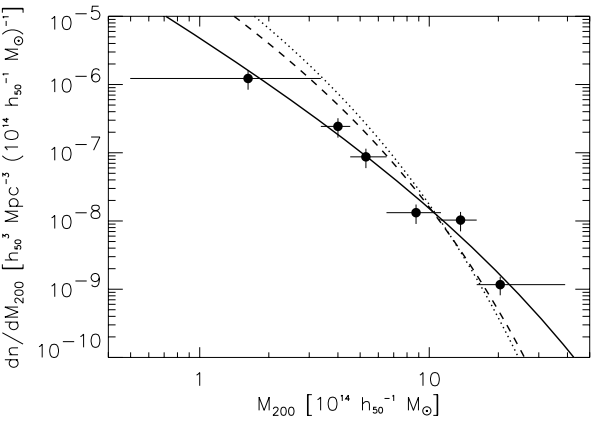


Fig. 4 Mass function of the HIFLUGCS X-ray clusters (*dots with error bars*). The *solid line* is the best fit with $\Omega_m = 0.12$ and $\sigma_8 = 0.98$. The *dashed* and *dotted lines* are the best fits with $\Omega_m = 0.5$ or $\Omega_m = 1.0$ held fixed, which yield $\sigma_8 = 0.60$ and $\sigma_8 = 0.46$, respectively. From Reiprich and Böhringer (2002)

In most mass estimation techniques, the assumption of dynamical equilibrium is fundamental for obtaining accurate estimates. There is however overwhelming evidence that many clusters are out of equilibrium, as described in the following sections.

2.2 Evidence of Non-equilibrium and Formation Processes

2.2.1 Cluster Galaxies

Elliptical and lenticulars (early-type) galaxies tend to reside in high-density regions, whereas spiral (late-type) galaxies are more common in low-density regions. This morphology-density relation has been known since a publication of Shapley in 1926, and possibly earlier. The first statistical analysis based on a large sample of 55 clusters showed that, in the local universe, the fraction of spirals is 80%, 60% and 0% in the field, outskirts and cores of clusters, respectively (Dressler 1980). Subsequent work confirmed the connection between galaxy properties and local density: spirals in clusters are H I- and dust-deficient, have larger metallicity and show an excess of radio continuum emission (see for a review Boselli and Gavazzi 2006).

At higher redshifts ($z \sim 1$), the fraction of spirals and irregulars in high-density regions increases at the expense of the fraction of lenticulars (Postman et al. 2005; Smith et al. 2005). Larger gas reservoirs and galaxy close encounters and collisions, which are more frequent in dense environments and favour the transformation of galaxy morphology, increase the star formation activity, as it occurs in starburst galaxies; indeed mid-infrared observations of luminous infrared galaxies show a more intense star formation activity in cluster environment at increasing redshift (Metcalfe et al. 2005; Bai et al. 2007). Although, on average, the properties of galaxies vary smoothly from the cluster centre to the outskirts, before reaching the average properties of galaxies in the field (e.g. Rines et al. 2005), the galaxy properties within individual clusters can show significant spatial variations, like the different slopes of the galaxy luminosity function in different regions of Coma (Adami et al. 2007).

In clusters with irregular morphology, the major axes of the bright galaxies tend to be aligned both with each other and with the major axis of the parent cluster (Plionis et al. 2003), as a century ago Wolf (1902) already pointed out for the galaxies in Coma. An alignment between the major axis of the bright central galaxy (BCG) or the cD galaxy and the major axis of the parent cluster is also present, as first noticed by Carter and Metcalfe

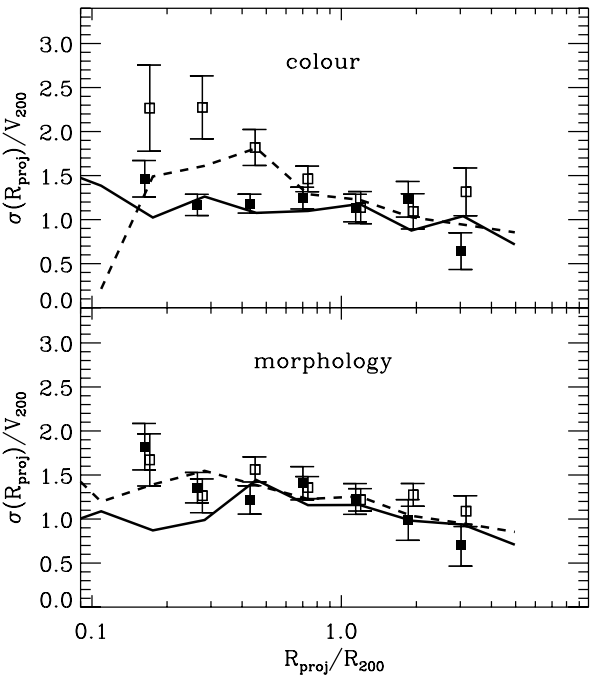


Fig. 5 Differential velocity dispersion profiles of galaxies brighter than $M_R = -20.5$ in the CNOC1 clusters ($0.18 < z < 0.55$). *Top panel*: Profiles of blue (*open squares*) and red (*filled squares*) galaxies; the rest-frame colour threshold is $B - V = 0.85$. The *dashed* and the *solid lines* show the corresponding profiles of the blue and red galaxies in mock clusters simulated with a N-body + semi-analytic technique (Diaferio et al. 2001; see Sect. 3). *Bottom panel*: Velocity dispersion profiles of disk-dominated galaxies (*open squares*) and bulge-dominated galaxies (*filled squares*). See Diaferio et al. (2001) for details

(1980). Any alignment present in the initial conditions is expected to be erased by multiple galaxy encounters occurring in the cluster environment. Therefore, the presence of alignment should indicate a young dynamical state of the cluster.

Blue cluster galaxies show larger velocity dispersions than red galaxies (Fig. 5; Biviano et al. 1997). This difference is attributed to more elongated orbits (Biviano and Katgert 2004), indicating that, unlike the red galaxies which are likely to be in dynamical equilibrium within the cluster gravitational potential well, blue galaxies are, on average, possibly falling onto the cluster for the first time.

Intracluster stars detected with planetary nebulae or as diffuse light in clusters at low (Gerhard et al. 2005) and intermediate redshift (Zibetti et al. 2005) can be the relics of tidal interactions suffered by falling galaxies (Covone et al. 2006) and/or the relics of merging of galaxies already in place in the forming gravitational potential well of the cluster (Murante et al. 2007), as observed in real clusters (Rines et al. 2007); intracluster stars can contain from 10 to 50% of the mass in stars in the core of clusters and can show distinct kinematic substructure (Arnaboldi et al. 2006).

2.2.2 Substructure, Mass Accretion and Non-thermal Phenomena

A large fraction of clusters shows the presence of substructure both in their galaxy distribution and in their X-ray emission morphology. The fraction of clusters with substructure depends on the cluster sample and on the substructure identification technique, but it is substantial, lying in the range 30–80% (see Ramella et al. 2007, and references therein). The X-ray band also shows patchy temperature (Belsole et al. 2005) and metallicity maps (e.g. Hayakawa et al. 2006; Finoguenov et al. 2006, Fig. 6; see Werner et al. 2008—Chap. 16, this issue). Moreover, the X-ray morphology is increasingly irregular with increasing redshift (Jeltema et al. 2005). Where the angular resolution is low and detailed maps of the

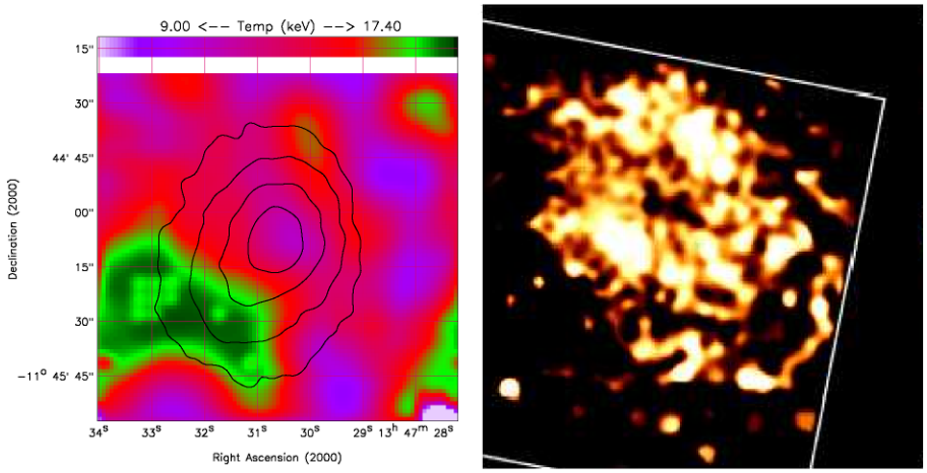


Fig. 6 *Left panel*: Temperature map of the cluster RX J1347.5-1145. The superposed contours show the X-ray surface brightness. See Gitti and Schindler (2004) for details. *Right panel*: Map of the metallicity abundance of the central region of A426, the Perseus cluster; brighter regions are more metal rich. See Sanders et al. (2005) for details

X-ray surface brightness cannot be obtained, a clear indication that, in some clusters, the ICM might be out of equilibrium comes from the hard, possibly thermal, excess (Rephaeli et al. 2008—Chap. 5, this issue), which appears when the X-ray spectrum is fitted assuming a plasma with single temperature and metallicity. In these cases, two or more temperatures are required to yield a reasonable fit.

All these pieces of evidence suggest that clusters accrete matter from the surrounding regions. The cluster 1ES 0657-55.8, dubbed "the bullet cluster" (Markevitch et al. 2002; Barrena et al. 2002), is a spectacular example of a merging cluster at intermediate redshift ($z = 0.296$). At low redshifts, there are many other examples of merging clusters or galaxy groups falling onto clusters along filamentary structure. To mention a few: Coma (Colless and Dunn 1996; Adami et al. 2005), A 521 (Ferrari et al. 2003), A 754 (Henry et al. 2004), A 2199 (Rines et al. 2001), A 2219 (Boschin et al. 2004), A 3560 in the Shapley supercluster (Bardelli et al. 2002), A 3921 (Ferrari et al. 2005). At high redshift the number of irregular and merging clusters increases substantially (Rosati 2004; Kodama et al. 2007).

Accretion of matter is a natural explanation for the patchy morphology of the X-ray emission and metallicity maps; it can also cause most of the turbulence in the ICM (Schuecker et al. 2004) and generate shocks (Bykov et al. 2008a—Chap. 7, this issue), which can be efficient generators of high-energy cosmic rays.

Non-thermal processes in the ICM appear in the X-ray and radio bands (Rephaeli et al. 2008—Chap. 5, this issue). Non-thermal hard X-ray emission is indeed observed in a few systems (e.g. Petrosian et al. 2006; Fusco-Femiano et al. 2007, and references therein). However, the existence of both the hard X-ray emission excess (Rossetti and Molendi 2004) and the soft X-ray and UV emission excesses, regardless of their thermal or non-thermal origin, are still debated (e.g. Lieu and Mittaz 2005; Nevalainen et al. 2007; Werner et al. 2007; Durret et al. 2008—Chap. 4, this issue).

At radio frequencies, the cluster emission can be both extended and associated with galaxies. One can infer the direction of the bulk motion of gas clouds within the ICM from

the deviations that the ICM kinetic pressure exerts on the radio jets emerging from active galactic nuclei (AGN) (Burns et al. 2002).

The extended radio emission has a steep spectrum and can appear with two different morphologies: either the emission comes from the cluster centre (halo), is extended and regular, or the emission comes from the cluster outskirts and is elongated (relic) (Feretti and Giovannini 2007; Ferrari et al. 2008—Chap. 6, this issue). By combined radio and hard X-ray observations or by Faraday rotation measurements one can infer the presence of magnetic fields in the ICM with intensity ranging from a few tenths to several μGauss (e.g. Blasi et al. 2007; Ferrari et al. 2008—Chap. 6, this issue).

Both the radio and the hard X-ray emissions originate from relativistic electrons: the former is synchrotron radiation, the latter is emission due to inverse Compton scattering with the CMB photons (Petrosian et al. 2008—Chap. 10, this issue). Among the many acceleration mechanisms proposed to produce the relativistic electrons (Petrosian and Bykov 2008—Chap. 11, this issue), shocks due to the accretion of matter from the surrounding regions seem to be one of the most efficient.

2.3 Connection with the Large-Scale Structure

2.3.1 A Morphology-Density Relation for Clusters

It has been known for a long time that clusters are not randomly distributed in the universe but form large concentrations named superclusters (Shapley 1933). The connection between clusters and the large-scale structure has become evident with extended galaxy redshift surveys, like the CfA (de Lapparent et al. 1986), the SDSS (York et al. 2000), and the 2dF (Colless et al. 2001) redshift surveys.

To quantify the distribution of objects in space, the simplest quantity to compute is the two-point correlation function $\xi(r)$, which yields the probability of finding two objects at separation r in excess to a Poisson distribution. Clusters correlate according to the correlation function $\xi = (r/r_0)^\gamma$, with $\gamma \approx -1.8$ (Bahcall and Soneira 1983). The correlation length r_0 is proportional to the cluster richness: richer clusters are rarer and their mean intercluster separation d larger. One finds $r_0 \propto d^{0.5}$, and $r_0 \approx 20$–$25h^{-1}$ Mpc for rich clusters. However, this $r_0 - d$ relation cannot be fully responsible for the large range, from ~ 10 to $\sim 30h^{-1}$ Mpc, covered by the correlation length r_0 when one considers clusters with increasing substructure (Fig. 7; Plionis and Basilakos 2002). Moreover, cluster shapes are generally elongated, as first quantified by Sastry (1968), and close clusters have their major axes aligned (Binggeli 1982). In the APM cluster sample, containing 903 objects, this alignment is stronger when the fraction of substructure is larger (Plionis and Basilakos 2002).

This connection between the properties of clusters and their large-scale environment is also apparent in the X-ray band: the X-ray luminosity of clusters (Böhringer 2004) and the irregularity of the X-ray surface brightness maps (Schuecker et al. 2001) are larger in environments with a higher cluster number density.

These features appear in a scenario (West 1994) where clusters form by episodic mass accretion along preferential directions (Colberg et al. 1999), in agreement with the currently accepted model of the formation of cosmic structure (see Sect. 3). Future estimates of cluster peculiar velocities based on the kinetic SZ effect, when corrected of their systematics (Diaferio et al. 2005), will further test whether the correlation between the evolution of the peculiar velocity and the local density predicted by this model (Diaferio et al. 2000; Sheth and Diaferio 2001) is correct.

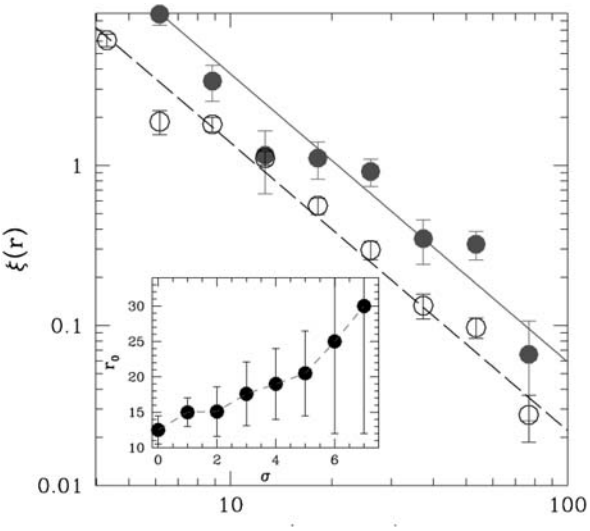


Fig. 7 Two-point correlation function of the 903 APM cluster sample (*open circles*) and of a subsample of clusters with significant substructure (*filled circles*). *Insert*: Correlation length r_0 vs. the significance level σ of substructure. See Plionis and Basilakos (2002) for details

2.3.2 The Surrounding WHIM

The optical depth of the intergalactic medium at $z \gtrsim 2$ derived from the Lyman-α forest of QSO spectra (Rauch et al. 1997), the abundances of light elements combined with the predictions of the standard Big Bang nucleosynthesis model and the power spectrum of the CMB (Spergel et al. 2007) are all consistent with a density of baryonic matter $\Omega_{\rm b}h^2 \simeq 0.022$. At low redshift, baryonic matter within galaxies and within the diffuse hot medium of galaxy clusters and groups only accounts for $\sim 15\%$ of this $\Omega_{\rm b}$ (Fukugita et al. 1998).

The hierarchical clustering model solves this discrepancy by predicting that 50% of the baryons permeating the matter distribution is in a warm phase with kT in the range 0.01–1 keV, overdensity in the range 0.1–10^4 and median $\sim$ 10–20, and median metallicity ~ 0.2 Solar, with a large spread, for oxygen (Cen and Ostriker 1999, 2006; Bertone et al. 2008—Chap. 14, this issue). The WHIM overdensity increases with increasing dark matter overdensity. Therefore, the outskirts of clusters and their environments are expected to be sources of a thermal soft X-ray emission. However, both the expected low surface brightness of the WHIM and the intergalactic and Galactic absorption make this observation very challenging. Nevertheless, there are claims of detection in many clusters, including Coma (Takei et al. 2007). In fact, rather than detecting the thermal X-ray emission (Durret et al. 2008—Chap. 4, this issue) directly, the most promising technique is searching for thermal UV and X-ray absorption lines in the spectra of bright background quasars or blazars (Richter et al. 2008—Chap. 3, this issue). Even with this strategy, current instruments (Paerels et al. 2008—Chap. 19, this issue) must be used at their sensitivity limits and the most plausible detections (e.g. Nicastro et al. 2005) are indeed still debated (Kaastra et al. 2006). Firm detections are however of great relevance, because the existence of the WHIM and the estimate of its properties would be an extraordinary confirmation of the current model of structure formation.

3 The Modelling Framework

3.1 Players: Gravity and Beyond

Modelling the formation and evolution of galaxy clusters requires the set-up of a complete cosmological context. According to the current paradigm of structure formation (as reviewed, e.g., in Springel et al. 2006), inflation amplifies quantum fluctuations, which are present in the dark matter density field of the early universe, to cosmic scales; at the time of matter and radiation density equality, these fluctuations start growing by gravitational instability. After radiation and ordinary baryonic matter decouple, radiation pressure stops supporting the baryonic density perturbations against self-gravity and against the pull of the dark matter gravitational potential wells that have already formed: the dark matter inhomogeneities thus accelerate the collapse of ordinary matter.

The pre-inflationary Gaussian quantum fluctuations generate a scale-free power spectrum of the primordial density perturbations. At later times, the dark matter particles, which have sufficient momenta to stream out of the denser regions, set in a characteristic scale in the power spectrum, which is proportional to the typical velocity of the dark matter particles themselves; below this scale the density perturbations are damped. The existence of quasars and galaxies at high redshift (Fan et al. 2004), the power spectrum of the CMB (Smoot et al. 1992; de Bernardis et al. 2000; Hanany et al. 2000), and the power spectrum of the large-scale distribution of galaxies (Tegmark et al. 2004) all agree with the hypothesis that the dark matter is "cold", namely that the dark matter particles become non-relativistic at very early times and the damping scale is $\ll 1h^{-1}$ Mpc. In these Cold Dark Matter (CDM) models, larger systems form by the aggregation of smaller systems and galaxy clusters are the nodes of the network of sheets and filaments which constitutes the matter density field (Fig. 8). Clusters can form at reasonably recent times, with the actual formation time depending on the cosmological parameters.

In CDM models, the scale-free properties of the primordial power spectrum are preserved on scales larger than the scales of galaxies. This feature implies that the dark matter halos of galaxies, groups and clusters evolve self-similarly. If gravitation is the only force playing,

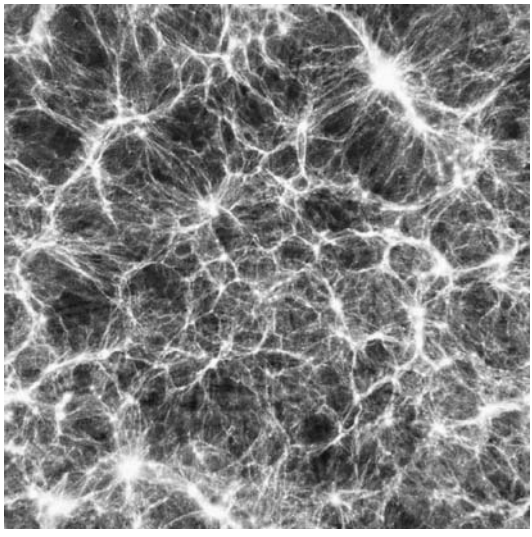


Fig. 8 Map density at $z = 0$ of the gas in a slice of the N-body/hydrodynamics simulation of Borgani et al. (2004). The volume shown is $192h^{-1}$ Mpc on a side and $12h^{-1}$ Mpc thick. The high-density regions are white, the low-density regions dark. See Borgani et al. (2004) for details

the physics of the ICM is ruled by adiabatic compression and shock heating, and we can derive simple relations between cluster properties at different redshifts (Kaiser 1986). The cluster virial mass scales as $M_{\rm vir} \propto \rho_{\rm c}(z)\Delta_{\rm c}(z)R^3$, where R is the cluster size in physical units, $\rho_{\rm c}(z) = 3H^2(z)/(8\pi G)$ is the critical density of the universe, and $\Delta_{\rm c}(z)$ is the cluster density in units of $\rho_{\rm c}(z)$; [2] $\rho_{\rm c}(z)$ scales with redshift z as $\rho_{\rm c}(z) \propto E^2(z) = \Omega_{\rm m}(1+z)^3 + (1-\Omega_{\rm m}-\Omega_\Lambda)(1+z)^2 + \Omega_\Lambda$. The cluster size thus scales as $R \propto M_{\rm vir}^{1/3}\Delta_{\rm c}^{-1/3}(z)E^{-2/3}(z)$, and the temperature as $T \propto M_{\rm vir}/R \propto M_{\rm vir}^{2/3}\Delta_{\rm c}^{1/3}(z)E^{2/3}(z)$; thus, for example, one finds for the X-ray luminosity due to Bremsstrahlung emission:

$$L_X \propto \int n_{\rm e} n_{\rm ions} \Lambda(T) {\rm d}V \propto M_{\rm vir}^2 R^{-3} T^{1/2} \propto M_{\rm vir}^{4/3} E^{7/3}(z) \Delta_{\rm c}^{7/6}(z) \tag{12}$$

or for the Comptonization parameter

$$y \propto \int n_{\rm e} T {\rm d}l \propto M_{\rm vir} R^{-2} T \propto M_{\rm vir} E^2(z) \Delta_{\rm c}(z). \tag{13}$$

These scaling relations only partially agree with observations; in fact, (1) the relative contribution of non-gravitational sources (AGN, galactic winds, supernovae, turbulence) to the ICM energy budget depends on the mass of the cluster and affects the slope of the scaling relations; and (2) the merging history of the cluster introduces an intrinsic scatter in the relations (Cavaliere et al. 1998). Moreover, clusters at high redshift have scaling relations steeper than the self-similar prediction (Ettori et al. 2004).

The deviations from the self-similar scaling relations are clearly related to non-gravitational processes. The interplay between gravitational and non-gravitational processes is highly non-linear and treating it appropriately requires sophisticated modelling. Over the last thirty years two main approaches have been followed: numerical simulations (Dolag et al. 2008a—Chap. 12, this issue) and an analytic (or semi-analytic) approach. Over the last ten years, the two approaches have been combined.

3.2 Making Clusters: Recipes, Ingredients and Tools

For the last thirty years, numerical simulations have represented the main tool of theoretical astrophysicists to investigate the non-linear evolution of cosmic structure and compare the theoretical results with observations. The first computational device for simulating galaxy interactions used an ingenious system of light-bulbs (Holmberg 1941). Numerical simulations of astrophysical systems have now reached tremendous sophistication (Dolag et al. 2008a—Chap. 12, this issue) when compared to the first numerical integration of the equations of motion of a few stars in a star cluster (von Hoerner 1960). Current numerical models reproduce the gross features of the large-scale structure, and the distribution of galaxies and diffuse baryons (Borgani et al. 2008a—Chap. 13; Bertone et al. 2008—Chap. 14, this issue), including the statistical properties of the Lyα forest (Springel et al. 2006).

Until the 1980s, numerical simulations relied on N-body techniques only to follow the evolution of the dissipationless component of cosmic structure. Particles in the simulations were first identified with galaxies and later on with dark matter particle tracers.

[2] A well-known approximation is

$$\Delta_{\rm c}(z) = 18\pi^2 + \begin{cases} 60w - 32w^2, & \Omega_{\rm m} \le 1,\ \Omega_\Lambda = 0, \\ 82w - 39w^2, & \Omega_{\rm m} + \Omega_\Lambda = 1, \end{cases} \tag{11}$$

where $w = \Omega_{\rm m}(z) - 1$ (Bryan and Norman 1998).

Peebles (1970) and White (1976) were the first to simulate the formation of a galaxy cluster with vacuum boundary conditions. To obtain a realistic distribution of matter, however, it is necessary to simulate a cosmological volume, with periodic boundary conditions, which is large enough to be representative of the universe. The first simulation of this kind, with a cold dark matter power spectrum of the initial density perturbations, demonstrated that dark matter must be constituted of cold collisionless particles, rather than massive neutrinos (Davis et al. 1985). Recent dark matter only simulations yield clusters whose number density evolution (Evrard et al. 2002), virial relations (Evrard et al. 2008), and connection between their shape and the large-scale structure (Altay et al. 2006; Lee and Evrard 2007) resemble the observations and can be used as probes of the cosmological model.

In the late 1980s, smoothed particle hydrodynamics (SPH) (Evrard 1988; Hernquist and Katz 1989) and mesh techniques (Cen et al. 1990) were introduced to follow the evolution of the dissipative component. The presence of dissipative gas substantially increases the number of physical processes to be taken into account and therefore increases the complexity of the simulations. In fact, *a priori*, it is not guaranteed that the different techniques yield similar results and some code comparisons have been performed over the years to check the consistency of the various integration schemes (Kang et al. 1994; Frenk et al. 1999; O'Shea et al. 2005).

The simplest approximation is to consider the gas non-radiative and perform adiabatic simulations (Evrard 1990). This model is able to reproduce the general trend of the self-similar scaling laws, but disagrees with observations in detail, yielding, in general, shallower X-ray luminosity-temperature and mass-temperature relations (Navarro et al. 1995b). To reproduce the observations more closely, and explain the deviations from the self-similar model, radiative cooling (Thomas and Couchman 1992; Katz and White 1993) and non-gravitational heating processes (Navarro et al. 1995b) must be included. More realistic models, attempting to reproduce the thermal properties of the ICM and the stellar content of clusters, have to include explicitly all those intimately coupled processes which rule galaxy formation and the thermodynamics of the ICM: star formation, energy feedback from supernovae explosion (Valdarnini 2002) and galactic winds (Metzler and Evrard 1994; Dolag et al. 2008a—Chap. 12; Borgani et al. 2008a—Chap. 13, this issue), besides metal production and chemical enrichment (Valdarnini 2003; Schindler and Diaferio 2008—Chap. 17; Borgani et al. 2008b—Chap. 18, this issue), which strongly affects the gas cooling rate; moreover, one needs to include magnetic fields and its associated non-thermal processes (Dolag et al. 1999, 2008b—Chap. 15, this issue).

The galaxy population in the simulated clusters have shown some rough similarities with observations since the very first numerical experiments (Evrard et al. 1994); many other observed properties are reproduced at different levels of agreement, including the soft X-ray thermal emission (Cheng et al. 2005). Nevertheless, in no simulation today the thermodynamical properties of the ICM, the stellar mass fraction and the galaxy luminosity function simultaneously match the observations. One of the most serious discrepancies with observations is the overproduction of stars: the fraction of baryons in stars usually is $\sim 20\%$ of the total baryon mass in clusters, at least a factor of two larger than observed (Borgani et al. 2004). A possible solution has been identified in the inhibition of the ICM cooling by the energy injection of AGN jets, as suggested by the X-ray cavities in correspondence of the radio emission observed around cluster central galaxies (McNamara et al. 2006; McNamara and Nulsen 2007); the presence of these cavities correlates with the X-ray temperature drops in the centre of some clusters, early interpreted as cooling flows (Dunn and Fabian 2006); however, the details of the heating mechanism by AGN jets remain uncertain (Voit 2005).

The main difficulty when treating most of the dissipative processes resides in the fact that they occur on scales much smaller than the spatial resolution of the simulation and must be approximated by phenomenological recipes: gas cooling, star formation and stellar evolution, which occur at scales $\ll 1$ pc, have effects on galactic and extragalactic scales $\gg 1$ Mpc; thus the simulations should model a spatial range of at least six orders of magnitude. These large dynamic ranges can be covered by neither a single numerical experiment nor experiments which rely on the resimulation technique, where an individual cluster extracted from a large-scale cosmological simulation is resimulated at a much higher resolution with appropriate boundary conditions (Navarro et al. 1995a).

The large number of physical processes and the large computational resources required to run an N-body/hydrodynamics simulation for a realistic modelling of cluster and galaxy formation have suggested a complementary strategy to explore more efficiently the parameter space: the combination of semi-analytic models with merger trees of dark matter halos extracted from dissipationless N-body simulations (Roukema et al. 1997; Kauffmann et al. 1999). This approach has reproduced the evolution and present-day spatial and kinematic distribution of galaxies in the universe and in clusters (Springel et al. 2005). The major shortcoming of this strategy is that severe uncertainties in the predictions can be introduced by the crude simplification of the physical processes and especially of the thermodynamics of the ICM, usually assumed to be isothermal and in hydrostatic equilibrium within each dark matter halo. Some improvements can be introduced by describing the ICM thermal properties in terms of entropy generation and distribution (McCarthy et al. 2007), but much work remains to be done.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". Partial support from the PRIN2006 grant "Costituenti fondamentali dell'Universo" of the Italian Ministry of University and Scientific Research and from the INFN grant PD51 is gratefully acknowledged. The authors also acknowledge financial support by the Austrian Science Foundation (FWF) through grants P18523-N16 and P19300-N16, by the Tiroler Wissenschaftsfonds and through the UniInfrastrukturprogramm 2005/06 by the BMWF.

References

G.O. Abell, Astrophys. J. Suppl. Ser. **3**, 211 (1958)
G.O. Abell, Annu. Rev. Astron. Astrophys. **3**, 1 (1965)
C. Adami, A. Biviano, F. Durret, A. Mazure, Astron. Astrophys. **443**, 17 (2005)
C. Adami, F. Durret, A. Mazure et al., Astron. Astrophys. **462**, 411 (2007)
G. Altay, J.M. Colberg, R.A.C. Croft, Mon. Not. R. Astron. Soc. **370**, 1422 (2006)
V.A. Ambartsumian, Astron. J. **66**, 536 (1961)
M. Arnaboldi, O. Gerhard, K.C. Freeman et al., in *Planetary Nebulae in Our Galaxy and Beyond*, ed. by M.J. Barlow, R.H. Méndez. Proc. IAU Symp. 234 (Cambridge Univ. Press, Cambridge, 2006), p. 337
N.A. Bahcall, R.M. Soneira, Astrophys. J. **270**, 20 (1983)
L. Bai, D. Marcillac, G.H. Rieke et al., Astrophys. J. **664**, 181 (2007)
S. Bardelli, T. Venturi, E. Zucca et al., Astron. Astrophys. **396**, 65 (2002)
R. Barrena, A. Biviano, M. Ramella, E.E. Falco, S. Seitz, Astron. Astrophys. **386**, 816 (2002)
J.G. Bartlett, astro-ph/0606241 (2006)
E. Belsole, J.-L. Sauvageot, G.W. Pratt, H. Bourdin, Adv. Space Res. **36**, 630 (2005)
S. Bertone, J. Schaye, K. Dolag, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9318-3
B. Binggeli, Astron. Astrophys. **107**, 338 (1982)
A. Biviano, Untangling Coma Berenices: A new vision of an old cluster, in *Proc. Meeting Held in Marseilles (France), June 17–20, 1997*, ed. by A. Mazure, F. Casoli, F. Durret, D. Gerbal (Word Scientific, Singapore, 1998), p. 1
A. Biviano, From Messier to Abell: 200 years of science with galaxy clusters, in *Constructing the Universe with Clusters of Galaxies*, ed. by F. Durret, D. Gerbal. IAP 2000 meeting, astro-ph/0010409 (2000)

A. Biviano, in *Proc. of the XLI Rencontres de Moriond*, ed. by L. Tresse, S. Maurogordato, J. Tran Thanh Van. XXVI Astrophysics Moriond Meeting: From dark halos to light, Editions Frontieres (2006)
A. Biviano, P. Katgert, Astron. Astrophys. **424**, 779 (2004)
A. Biviano, P. Katgert, A. Mazure et al., Astron. Astrophys. **321**, 84 (1997)
P. Blasi, S. Gabici, G. Brunetti, Int. J. Mod. Phys. **A22**, 681 (2007)
H. Böhringer, The large-scale environment of groups and clusters of galaxies, in *Outskirts of Galaxy Clusters: Intense Life in the Suburbs*, ed. by A. Diaferio. IAU Coll. 195 (2004), p. 12
M. Bonamente, M.K. Joy, S.J. LaRoque et al., Astrophys. J. **647**, 25 (2006)
S. Borgani, G. Murante, V. Springel et al., Mon. Not. R. Astron. Soc. **348**, 1078 (2004)
S. Borgani, A. Diaferio, K. Dolag, S. Schindler, Space Sci. Rev. (2008a). doi:10.1007/s11214-008-9317-4
S. Borgani, D. Fabjan, L. Tornatore et al., Space Sci. Rev. (2008b). doi:10.1007/s11214-008-9322-7
W. Boschin, M. Girardi, R. Barrena et al., Astron. Astrophys. **416**, 839 (2004)
A. Boselli, G. Gavazzi, Publ. Astron. Soc. Pac. **118**, 517 (2006)
G.L. Bryan, M.L. Norman, Astrophys. J. **495**, 80 (1998)
J.O. Burns, C. Loken, K. Roettiger et al., New Astron. Rev. **46**, 135 (2002)
A.M. Bykov, K. Dolag, F. Durret, Space Sci. Rev. (2008a). doi:10.1007/s11214-008-9312-9
A.M. Bykov, F.B.S. Paerels, V. Petrosian, Space Sci. Rev. (2008b). doi:10.1007/s11214-008-9309-4
J.E. Carlstrom, G.P. Holder, E.D. Reese, Annu. Rev. Astron. Astrophys. **40**, 643 (2002)
D. Carter, N. Metcalfe, Mon. Not. R. Astron. Soc. **191**, 325 (1980)
A. Cavaliere, N. Menci, P. Tozzi, Astrophys. J. **501**, 493 (1998)
R. Cen, J.P. Ostriker, Astrophys. J. **514**, 1 (1999)
R. Cen, J.P. Ostriker, Astrophys. J. **650**, 560 (2006)
R.Y. Cen, J.P. Ostriker, A. Jameson, F. Liu, Astrophys. J. **362**, L41 (1990)
L.-M. Cheng, S. Borgani, P. Tozzi et al., Astron. Astrophys. **431**, 405 (2005)
J.M. Colberg, S.D.M. White, A. Jenkins, F.R. Pearce, Mon. Not. R. Astron. Soc. **308**, 593 (1999)
M. Colless, A.M. Dunn, Astrophys. J. **458**, 435 (1996)
M. Colless, G. Dalton, S. Maddox et al., Mon. Not. R. Astron. Soc. **328**, 1039 (2001)
G. Covone, C. Adami, F. Durret et al., Astron. Astrophys. **460**, 381 (2006)
L.L. Cowie, M. Henriksen, R. Mushotzky, Astrophys. J. **317**, 593 (1987)
H. Dahle, astro-ph/0701598 (2007)
M. Davis, G. Efstathiou, C.S. Frenk, S.D.M. White, Astrophys. J. **292**, 371 (1985)
P. de Bernardis, P.A.R. Ade, J.J. Bock et al., Nature **404**, 955 (2000)
V. de Lapparent, M.J. Geller, J.P. Huchra, Astrophys. J. **302**, L1 (1986)
A. Diaferio, Mon. Not. R. Astron. Soc. **309**, 610 (1999)
A. Diaferio, M.J. Geller, Astrophys. J. **481**, 633 (1997)
A. Diaferio, R.A. Sunyaev, A. Nusser, Astrophys. J. **533**, L71 (2000)
A. Diaferio, G. Kauffmann, M.L. Balogh et al., Mon. Not. R. Astron. Soc. **323**, 999 (2001)
A. Diaferio, S. Borgani, L. Moscardini et al., Mon. Not. R. Astron. Soc. **356**, 1477 (2005)
K. Dolag, M. Bartelmann, H. Lesch, Astron. Astrophys. **348**, 351 (1999)
K. Dolag, S. Borgani, S. Schindler, A. Diaferio, A.M. Bykov, Space Sci. Rev. (2008a). doi:10.1007/s11214-008-9316-5
K. Dolag, A.M. Bykov, A. Diaferio, Space Sci. Rev. (2008b). doi:10.1007/s11214-008-9319-2
A. Dressler, Astrophys. J. **236**, 351 (1980)
R.J.H. Dunn, A.C. Fabian, Mon. Not. R. Astron. Soc. **373**, 959 (2006)
F. Durret, J.S. Kaastra, J. Nevalainen, T. Ohashi, N. Werner, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9313-8
S. Ettori, P. Tozzi, S. Borgani, P. Rosati, Astron. Astrophys. **417**, 13 (2004)
A.E. Evrard, Mon. Not. R. Astron. Soc. **235**, 911 (1988)
A.E. Evrard, Astrophys. J. **363**, 349 (1990)
A.E. Evrard, F.J. Summers, M. Davis, Astrophys. J. **422**, 11 (1994)
A.E. Evrard, T.J. MacFarland, H.M.P. Couchman et al., Astrophys. J. **573**, 7 (2002)
A.E. Evrard, J. Bialek, M. Busha et al., Astrophys. J. **672**, 122 (2008)
X. Fan, J.F. Hennawi, G.T. Richards et al., Astron. J. **128**, 515 (2004)
J.E. Felten, R.J. Gould, W.A. Stein, N.J. Woolf, Astrophys. J. **146**, 955 (1966)
L. Feretti, G. Giovannini, in *Panchromatic View of Clusters of Galaxies and the Large-Scale Structure*, ed. by M. Plionis, O. Lopez-Cruz, D. Hughes. Springer Lect. Notes in Phys. (Springer, Berlin, 2007 in press). (astro-ph/0703494)
C. Ferrari, S. Maurogordato, A. Cappi, C. Benoist, Astron. Astrophys. **399**, 813 (2003)
C. Ferrari, F. Govoni, S. Schindler, A. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9311-x
C. Ferrari, C. Benoist, S. Maurogordato, A. Cappi, E. Slezak, Astron. Astrophys. **430**, 19 (2005)

A. Finoguenov, M.J. Henriksen, F. Miniati, U.G. Briel, C. Jones, Astrophys. J. **643**, 790 (2006)
C.S. Frenk, S.D.M. White, P. Bode et al., Astrophys. J. **525**, 554 (1999)
M. Fukugita, C.J. Hogan, P.J.E. Peebles, Astrophys. J. **503**, 518 (1998)
R. Fusco-Femiano, R. Landi, M. Orlandini, Astrophys. J. **654**, L9 (2007)
O. Gerhard, M. Arnaboldi, K.C. Freeman et al., Astrophys. J. **621**, L93 (2005)
M. Girardi, G. Giuricin, F. Mardirossian, M. Mezzetti, W. Boschin, Astrophys. J. **505**, 74 (1998)
M. Gitti, S. Schindler, Astron. Astrophys. **427**, L9 (2004)
H. Gursky, R. Levinson, E. Kellogg et al., Astrophys. J. **173**, L99 (1972)
S. Hanany, P. Ade, A. Balbi et al., Astrophys. J. **545**, L5 (2000)
A. Hayakawa, A. Hoshino, M. Ishida et al., Publ. Astron. Soc. Jpn. **58**, 695 (2006)
J.F. Hennawi, M.D. Gladders, M. Oguri et al., Astron. J. **135**, 664 (2008)
J.P. Henry, A. Finoguenov, U.G. Briel, Astrophys. J. **615**, 181 (2004)
L. Hernquist, N. Katz, Astrophys. J. Suppl. Ser. **70**, 419 (1989)
E. Holmberg, Astrophys. J. **94**, 385 (1941)
T.E. Jeltema, C.R. Canizares, M.W. Bautz, D.A. Buote, Astrophys. J. **624**, 606 (2005)
J.S. Kaastra, N. Werner, J.W.A. den Herder et al., Astrophys. J. **652**, 189 (2006)
J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-y
N. Kaiser, Mon. Not. R. Astron. Soc. **222**, 323 (1986)
H. Kang, J.P. Ostriker, R. Cen et al., Astrophys. J. **430**, 83 (1994)
N. Katz, S.D.M. White, Astrophys. J. **412**, 455 (1993)
G. Kauffmann, J.M. Colberg, A. Diaferio, S.D.M. White, Mon. Not. R. Astron. Soc. **303**, 188 (1999)
T. Kodama, M. Tanaka, I. Tanaka, M. Kajisawa, in *Galaxy Evolution Across the Hubble Time*, ed. by F. Combes, J. Palous. IAU Symp. 235 (2007), p. 170
J. Lee, A.E. Evrard, Astrophys. J. **657**, 30 (2007)
R. Lieu, J.P.D. Mittaz, The Cluster soft excess: new faces of an old enigma, in *The Identification of Dark Matter*, ed. by N.J.C. Spooner, V. Kudryavtsev (World Scientific, Singapore, 2005), p. 18
R. Lynds, V. Petrosian, Bull. Am. Astron. Soc. **18**, 1014 (1986)
M. Markevitch, A.H. Gonzalez, L. David et al., Astrophys. J. **567**, L27 (2002)
I.G. McCarthy, R.G. Bower, M.L. Balogh et al., Mon. Not. R. Astron. Soc. **376**, 497 (2007)
B.R. McNamara, P.E.J. Nulsen, Annu. Rev. Astron. Astrophys. **45**, 117 (2007)
B.R. McNamara, D.A. Rafferty, L. Bîrzan et al., Astrophys. J. **648**, 164 (2006)
L. Metcalfe, D. Fadda, A. Biviano, Space Sci. Rev. **119**, 425 (2005)
C.A. Metzler, A.E. Evrard, Astrophys. J. **437**, 564 (1994)
S. Muchovej, T. Mroczkowski, J.E. Carlstrom et al., Astrophys. J. **663**, 708 (2007)
G. Murante, M. Giovalli, O. Gerhard et al., Mon. Not. R. Astron. Soc. **377**, 2 (2007)
J.F. Navarro, C.S. Frenk, S.D.M. White, Mon. Not. R. Astron. Soc. **275**, 56 (1995a)
J.F. Navarro, C.S. Frenk, S.D.M. White, Mon. Not. R. Astron. Soc. **275**, 720 (1995b)
J. Nevalainen, M. Bonamente, J. Kaastra, Astrophys. J. **656**, 733 (2007)
F. Nicastro, S. Mathur, M. Elvis et al., Nature **433**, 495 (2005)
B.W. O'Shea, K. Nagamine, V. Springel, L. Hernquist, M.L. Norman, Astrophys. J. Suppl. Ser. **160**, 1 (2005)
F. Paerels, J.S. Kaastra, T. Ohashi et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9323-6
P.J.E. Peebles, Astron. J. **75**, 13 (1970)
V. Petrosian, G. Madejski, K. Luli, Astrophys. J. **652**, 948 (2006)
V. Petrosian, A.M. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9315-6
V. Petrosian, A.M. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9327-2
M. Plionis, S. Basilakos, Mon. Not. R. Astron. Soc. **329**, L47 (2002)
M. Plionis, C. Benoist, S. Maurogordato, C. Ferrari, S. Basilakos, Astrophys. J. **594**, 144 (2003)
M. Postman, M. Franx, N.J.G. Cross et al., Astrophys. J. **623**, 721 (2005)
W.H. Press, P. Schechter, Astrophys. J. **187**, 425 (1974)
M. Ramella, A. Biviano, A. Pisani et al., Astron. Astrophys. **470**, 39 (2007)
M. Rauch, J. Miralda-Escudé, W.L.W. Sargent et al., Astrophys. J. **489**, 7 (1997)
T.H. Reiprich, H. Böhringer, Astrophys. J. **567**, 716 (2002)
Y. Rephaeli, J. Nevalainen, T. Ohashi, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9314-7
P. Richter, F. Paerels, J.S. Kaastra, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9325-4
K. Rines, A. Mahdavi, M.J. Geller et al., Astrophys. J. **555**, 558 (2001)
K. Rines, M.J. Geller, M.J. Kurtz, A. Diaferio, Astron. J. **130**, 1482 (2005)
K. Rines, R. Finn, A. Vikhlinin, Astrophys. J. **665**, L9 (2007)
P. Rosati, in *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, ed. by J.S. Mulchaey, A. Dressler, A. Oemler. Carnegie Obs. Astrophys. Ser. (Cambridge Univ. Press, Cambridge, 2004), p. 72

P. Rosati, S. Borgani, C. Norman, Annu. Rev. Astron. Astrophys. **40**, 539 (2002)
M. Rossetti, S. Molendi, Astron. Astrophys. **414**, L41 (2004)
B.F. Roukema, P.J. Quinn, B.A. Peterson, B. Rocca-Volmerange, Mon. Not. R. Astron. Soc. **292**, 835 (1997)
J.S. Sanders, A.C. Fabian, R.J.H. Dunn, Mon. Not. R. Astron. Soc. **360**, 133 (2005)
D.N. Sastry, Publ. Astron. Soc. Pac. **80**, 252 (1968)
S. Schindler, A. Diaferio, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9321-8
P. Schneider, in *Gravitational Lensing: Strong, Weak and Micro*, ed. by G. Meylan, P. Jetzer, P. North. Saas-Fee Advanced Course, vol. 33 (Springer, Berlin, 2006), p. 1
P. Schuecker, H. Böhringer, T.H. Reiprich, L. Feretti, Astron. Astrophys. **378**, 408 (2001)
P. Schuecker, A. Finoguenov, F. Miniati, H. Böhringer, U.G. Briel, Astron. Astrophys. **426**, 387 (2004)
H. Shapley, Harvard. Obs. Bull. **838**, 3 (1926)
H. Shapley, Proc. Nat. Acad. Sci. Washington **19**, 591 (1933)
R.K. Sheth, A. Diaferio, Mon. Not. R. Astron. Soc. **322**, 901 (2001)
R.K. Sheth, G. Tormen, Mon. Not. R. Astron. Soc. **308**, 119 (1999)
G.P. Smith, T. Treu, R.S. Ellis, S.M. Moran, A. Dressler, Astrophys. J. **620**, 78 (2005)
G.F. Smoot, C.L. Bennet, A. Kogut et al., Astrophys. J. **396**, L1 (1992)
D.N. Spergel, R. Bean, O. Doré et al., Astrophys. J. Suppl. Ser. **170**, 377 (2007)
V. Springel, S.D.M. White, A. Jenkins et al., Nature **435**, 629 (2005)
V. Springel, C.S. Frenk, S.D.M. White, Nature **440**, 1137 (2006)
R.A. Sunyaev, Ya.B. Zeldovich, Comments Astrophys. Space Phys. **4**, 173 (1972)
Y. Takei, J.P. Henry, A. Finoguenov et al., Astrophys. J. **655**, 831 (2007)
M. Tegmark, M.R. Blanton, M.A. Strauss et al., Astrophys. J. **606**, 702 (2004)
L.S. The, S.D.M. White, Astron. J. **92**, 1248 (1986)
P.A. Thomas, H.M.P. Couchman, Mon. Not. R. Astron. Soc. **257**, 11 (1992)
R. Valdarnini, Astrophys. J. **567**, 741 (2002)
R. Valdarnini, Mon. Not. R. Astron. Soc. **339**, 1117 (2003)
G.M. Voit, Rev. Mod. Phys. **77**, 207 (2005)
S. von Hoerner, Z. Astrophys. **50**, 184 (1960)
N. Werner, J.S. Kaastra, Y. Takei et al., Astron. Astrophys. **468**, 849 (2007)
N. Werner, F. Durret, T. Ohashi, S. Schindler, R.P.C. Wiersma, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9320-9
M.J. West, Mon. Not. R. Astron. Soc. **268**, 79 (1994)
S.D.M. White, Mon. Not. R. Astron. Soc. **177**, 717 (1976)
M. Wolf, Publ. Astrophys. Inst. Königstuhl-Heidelberg **1**, 125 (1902)
D.G. York, J. Adelman, J.E. Anderson Jr. et al., Astron. J. **120**, 1579 (2000)
S. Zibetti, S.D.M. White, D.P. Schneider, J. Brinkmann, Mon. Not. R. Astron. Soc. **358**, 949 (2005)
F. Zwicky, Helv. Phys. Acta **6**, 110 (1933)
F. Zwicky, Astrophys. J. **86**, 217 (1937)
F. Zwicky, E. Herzog, P. Wild, M. Karpowicz, C.T. Kowal, *Catalogue of Galaxies and Clusters of Galaxies* (Pasadena, California Institute of Technology, 1961–1968)

Chapter 3
FUV and X-Ray Absorption in the Warm-Hot Intergalactic Medium

P. Richter · F.B.S. Paerels · J.S. Kaastra

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9325-4

Abstract The Warm-Hot Intergalactic Medium (WHIM) arises from shock-heated gas collapsing in large-scale filaments and probably harbours a substantial fraction of the baryons in the local Universe. Absorption-line measurements in the ultraviolet (UV) and in the X-ray band currently represent the best method to study the WHIM at low redshifts. We here describe the physical properties of the WHIM and the concepts behind WHIM absorption line measurements of H I and high ions such as O VI, O VII, and O VIII in the far-ultraviolet and X-ray band. We review results of recent WHIM absorption line studies carried out with UV and X-ray satellites such as FUSE, HST, Chandra, and XMM-Newton and discuss their implications for our knowledge of the WHIM.

Keywords Galaxies: intergalactic medium · Quasars: absorption lines · Cosmology: large-scale structure of the Universe

1 Introduction

As recent cosmological simulations imply, the temperature of the intergalactic medium (IGM) undergoes a significant change from high to low redshifts parallel to the proceeding of large-scale structure formation in the Universe (e.g., Cen and Ostriker 1999; Davé et al. 2001). As a result, a substantial fraction of the baryonic matter in the local Universe is expected to reside in the so-called Warm-Hot Intergalactic Medium (WHIM). The WHIM represents a low-density ($n_{\rm H} \sim 10^{-6}$–10^{-4} cm^{-3}), high-temperature ($T \sim 10^5$–10^7 K) plasma

P. Richter (✉)
Institut für Physik, Universität Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany
e-mail: prichter@astro.physik.uni-potsdam.de

F.B.S. Paerels
Department of Astronomy and Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA

J.S. Kaastra
SRON, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands

that primarily is made of protons, electrons, He II, and He III, together with traces of some highly-ionised heavy elements. The WHIM is believed to emerge from intergalactic gas that is shock-heated to high temperatures as the medium is collapsing under the action of gravity in large-scale filaments (e.g., Valageas et al. 2002). In this scenario, part of the warm (photoionised) intergalactic medium that gives rise to the Ly α forest in the spectra of distant quasars (QSO) is falling into the potential wells of the increasingly pronounced filaments, gains energy (through gravity), and is heated to high temperatures by shocks that run through the plasma.

Because of the low density and the high degree of ionisation, direct observations of the shock-heated and collisionally ionised WHIM are challenging with current instrumentation (in contrast to the photoionised IGM, which is easily observable through the Ly α forest). Diffuse emission from the WHIM plasma must have a very low surface brightness and its detection awaits UV and X-ray observatories more sensitive than currently available (see, e.g., Fang et al. 2005; Kawahara et al. 2006). The most promising approach to study the WHIM with observations at low redshift is to search for absorption features from the WHIM in FUV and in the X-ray regime in the spectra of quasars, active galactic nuclei (AGN) and other suited extragalactic background sources. As the WHIM represents a highly-ionised plasma, the most important WHIM absorption lines are those originating from the electronic transitions of high-ionisation state ions (hereafter referred to as "high ions") of abundant heavy elements such as oxygen and carbon. Among these, five-times ionised oxygen (O VI) is the most valuable high ion to trace the WHIM at temperatures of $T \sim 3 \times 10^5$ K in the FUV regime. In the X-ray band, the O VII and O VIII transitions represent the key observables to trace the WHIM at higher temperatures in the range $3 \times 10^5 < T < 10^7$. In addition to the spectral signatures of high ions of heavy elements the search for broad and shallow Ly α absorption from the tiny fraction of neutral hydrogen in the WHIM represents another possibility to identify and study the most massive WHIM filaments in the intergalactic medium with FUV absorption spectroscopy. Finally, for the interpretation of the observed WHIM absorption features in UV and X-ray spectra the comparison between real data and artificial spectra generated by numerical simulations that include realistic gas physics is of great importance to identify possible pitfalls related to technical and physical issues such as limited signal-to-noise ratios and spectral resolution, line-broadening mechanisms, non-equilibrium conditions, and others.

In this chapter, we review the physics and methodology of the UV and X-ray absorption measurements of warm-hot intergalactic gas at low redshift and summarise the results of recent observations obtained with space-based observatories. The outline of this chapter is the following. The ionisation conditions of the WHIM and the most important absorption signatures of this gas in the UV and X-ray band are presented in Sect. 2. Recent UV absorption measurements of the WHIM at low redshift are discussed in Sect. 3. Similarly, measurements of the WHIM in the X-ray are presented in Sect. 4. In Sect. 5 we compare the results from WHIM observations with predictions from numerical simulations and give an overview of WHIM measurements at high redshift. Finally, some concluding remarks are given in Sect. 6.

2 Physical Properties of the WHIM

2.1 WHIM Ionisation Conditions

The occurrence and characteristics of the WHIM absorption signatures in the FUV and X-ray band are determined to a high degree by the ionisation conditions in the gas. We briefly

discuss the WHIM ionisation properties, as this is crucial for interpretation of the WHIM absorption lines in FUV and X-ray spectra that arise in such warm-hot gas. Generally, there are two processes that determine the ionisation state of warm-hot gas in the intergalactic medium: collisional ionisation caused by the high temperature of the gas in collapsed structures and photoionisation by the cosmic FUV background.

2.1.1 Hydrogen

By far most of the mass of the WHIM is in the form of ionised hydrogen. Therefore, understanding the processes that lead to the ionisation of hydrogen is essential for the interpretation of WHIM absorption lines and for a reliable estimate of the baryon content of warm-hot intergalactic gas. The ionisation potential of neutral hydrogen is 13.6 eV and thus both ionisation by particle collisions and ionisation by high-energy photons contribute to the ionisation of H I in warm-hot gas. We start with collisional ionisation, which is believed to dominate the ionisation of hydrogen at temperatures $> 10^5$ K.

In collisional ionisation equilibrium (CIE)—the most simple approach to characterise the ionisation conditions in low-density, high-temperature plasmas—the ionisation fraction depends only on the gas temperature. If we ignore any charge-exchange reactions (which is justified in case of hydrogen), the neutral hydrogen fraction in CIE is simply the ratio between the recombination coefficient $\alpha_{\rm H}(T)$ and the collisional ionisation coefficient $\beta_{\rm H}(T)$:

$$f_{\rm H\,I,coll} = \frac{\alpha_{\rm H}(T)}{\beta_{\rm H}(T)}. \tag{1}$$

Above gas temperatures of $\sim 1.5 \times 10^4$ K collisions by thermal electrons efficiently ionise hydrogen to a high degree, and already at $T \sim 3 \times 10^4$ K the neutral hydrogen fraction in the gas is less then one percent. For the temperature range that is characteristic for the WHIM, $T = 10^5 - 10^7$ K, one can approximate the ionisation fraction in a collisional ionisation equilibrium in the way

$$\log f_{\rm H\,I,coll} \approx 13.9 - 5.4 \log T + 0.33\,(\log T)^2. \tag{2}$$

where T is in units K (Richter et al. 2006a; Sutherland and Dopita 1993). Thus, for WHIM gas with $T = 10^6$ K the neutral hydrogen fraction in the gas in CIE is only $\sim 2.4 \times 10^{-7}$.

Next to particle collisions, photons with energies > 13.6 eV contribute to the ionisation of the WHIM, in particular in the low-temperature WHIM tail at $\sim 10^5$ K and below. Such ionising photons in intergalactic space are indeed provided by the metagalactic ultraviolet (UV) background, originating from the hard radiation emitted by QSOs and AGN. Figure 1 shows the spectral shape of the UV background at $z = 0$ (left panel) and the redshift-dependence of the hydrogen photoionisation rate from the UV background (right panel) based on the models by Haardt and Madau (1996).

Considering photoionisation, one generally can write for the neutral-hydrogen fraction in the gas:

$$f_{\rm H\,I,photo} = \frac{n_{\rm e}\,\alpha_{\rm H}(T)}{\Gamma_{\rm H\,I}}, \tag{3}$$

where $\alpha_{\rm H}(T)$ denotes the temperature-dependent recombination rate of hydrogen, $n_{\rm e}$ is the electron density, and $\Gamma_{\rm H\,I}$ is the photoionisation rate. $\Gamma_{\rm H\,I}$ depends on the ambient ionising

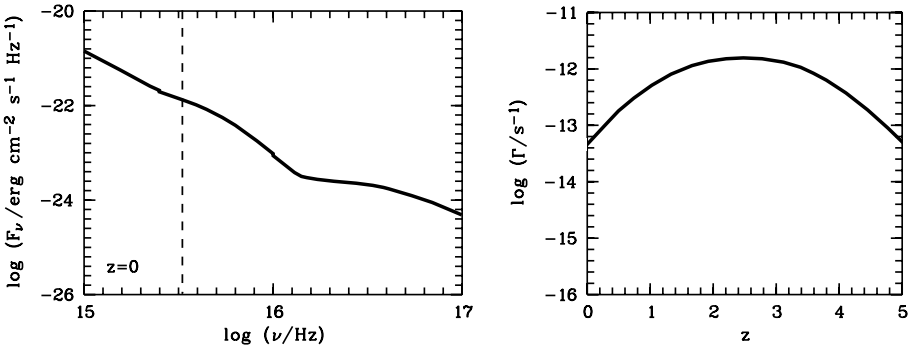


Fig. 1 *Left panel:* Spectral shape of the metagalactic UV background at $z = 0$ (from Haardt and Madau 1996). Plotted is the flux of photons ($F_\nu = 4\pi J_\nu$) against the frequency ν. The hydrogen ionisation edge is indicated with a dashed line. *Right panel:* Redshift-dependence of the hydrogen photoionisation rate Γ from the UV background for the range $z = 0$ to $z = 5$. Adapted from Haardt and Madau (1996)

radiation field J_ν (in units erg cm^{-2} s^{-1} Hz^{-1} sr^{-1}) in the WHIM provided by the metagalactic UV background (see Fig. 1):

$$\Gamma_{\mathrm{H\,I}} = 4\pi \int_{\nu_\mathrm{L}}^{\infty} \frac{\sigma_\nu J_\nu}{\mathrm{h}\nu}\,\mathrm{d}\nu \approx 2.5 \times 10^{-14}\, J_{-23}\ \mathrm{s}^{-1}. \tag{4}$$

Here, ν_L is the frequency at the Lyman limit and σ_ν denotes the photoionisation cross section of hydrogen, which scales with ν^{-3} for frequencies larger than ν_L (see Kaastra et al. 2008—Chap 9, this issue). We have introduced the dimensionless scaling factor J_{-23} which gives the metagalactic UV radiation intensity at the Lyman limit in units 10^{-23} erg cm^{-2} s^{-1} Hz^{-1} sr^{-1}. For $z = 0$ we have $J_{-23} \sim 1{-}2$, while for $z = 3$ the value for J_{-23} is ~ 80, thus significantly higher (Haardt and Madau 1996). Assuming $n_\mathrm{e} = n_\mathrm{H}$ and inserting a proper function for $\alpha_\mathrm{H}(T)$, we finally can write for the logarithmic neutral hydrogen fraction in a purely photoionised WHIM plasma

$$\log f_{\mathrm{H\,I,photo}} \approx \log\left(\frac{16\, n_\mathrm{H}\, T_4^{-0.76}}{J_{-23}}\right), \tag{5}$$

where n_H is the hydrogen volume density in units cm^{-3} and T_4 is the temperature in units 10^4 K. Thus, for purely photoionised intergalactic gas at $z = 0$ with $n = 5 \times 10^{-6}$ and $T = 10^6$ K we find that the neutral hydrogen fraction is $f_{\mathrm{H\,I,photo}} \sim 2.4 \times 10^{-6}$. This is ten times higher than for CIE, indicating that collisions dominate the ionisation fraction of hydrogen in intermediate and high-temperature WHIM regions. However, note that at lower temperatures near $T = 10^5$ K at the same density we have $f_{\mathrm{H\,I,photo}} \sim f_{\mathrm{H\,I,coll}}$. Since this is the WHIM temperature regime preferentially detected by UV absorption features (e.g., O VI and broad Ly α), photoionisation is important and needs to be accounted for when it comes to the interpretation of WHIM absorbers observed in the FUV. From a WHIM simulation at $z = 0$ including both collisional ionisation and photoionisation Richter et al. (2006b; see Fig. 2) find the following empirical relation between neutral hydrogen fraction and gas temperature for a WHIM density range between $\log n_\mathrm{H} = -5.3$ and -5.6:

$$\log f_{\mathrm{H\,I}} \approx 0.75 - 1.25 \log T. \tag{6}$$

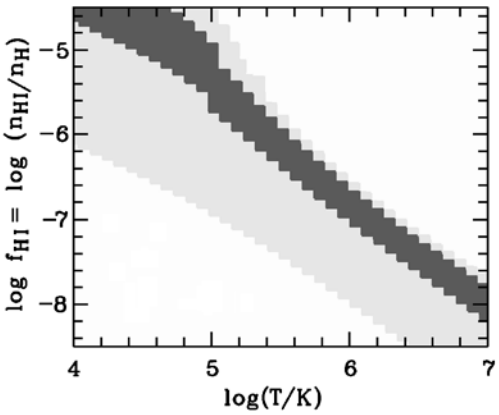


Fig. 2 The neutral hydrogen fraction, $\log f_{\rm HI} = \log(n_{\rm HI}/n_{\rm H})$, in a WHIM simulation (photoionisation+collisional ionisation), is plotted as a function of the gas temperature, $\log T$. The *light gray shaded* indicates cells in the density range $\log n_{\rm H} = -5$ to -7. The *dark gray shaded area* refers to cells that have $\log n_{\rm H} = -5.3$ to -5.6, thus a density range that is characteristic for WHIM absorbers. Adapted from Richter et al. (2006b)

This equation may serve as a thumb rule to estimate ionisation fractions in WHIM absorbers at $z = 0$ if the gas temperature is known (e.g., from measurements of the line widths; see Sect. 2.2.1).

2.1.2 Oxygen and Other Metals

While hydrogen provides most of the mass in the WHIM, the most important diagnostic lines to study this gas phase are from highly ionised metals such as oxygen, neon, carbon, magnesium, and others. Therefore, the understanding of the ionisation properties of the observed high ions of these elements is as important as for hydrogen. As for hydrogen, both collisional ionisation and photoionisation need to be considered. With its single electron, hydrogen can only be either neutral or fully ionised. Heavy elements, in contrast, have several electrons available and are—even at very high temperatures—usually only partly ionised. Thus, electronic transitions exist for such highly-ionised metals ("high ions") in warm-hot gas. Of particular importance for observations of the WHIM are the high ionisation states of oxygen, O VI, O VII, and O VIII, as they have strong electronic transitions in the UV (O VI) and at X-ray wavelengths (O VII & O VIII) and oxygen is a relatively abundant element. Another important metal for observing warm-hot gas in the UV and X-ray band is neon (Ne VII, Ne VIII, Ne IX, Ne X). In collisional ionisation equilibrium, the ionisation state of these elements is determined solely by the temperature of the gas. For each element, the ionisation fractions of the ionisation states (e.g., four-times vs. five-times ionised) then are characterised by the respective ionisation potentials (IPs) of the individual ionisation levels. For instance, at $T \sim 1{-}3 \times 10^5$ K, a significant fraction of the oxygen is five-times ionised (O^{+5} or O VI, IP = 138 eV). Six-times ionised oxygen (O^{+6} or O VII, IP = 739 eV) and seven-times ionised oxygen (O^{+7} or O VIII, IP = 871 eV) predominantly exist at higher temperatures in the range $3 \times 10^5{-}3 \times 10^6$ K and $3 \times 10^6{-}10^7$ K, respectively. Figure 3 shows the ionisation fractions of the most important high ions of oxygen and neon, based on the CIE calculations of Sutherland and Dopita (1993); see also Kaastra et al. (2008—Chap. 9, this issue). High ions of other elements such as carbon, nitrogen, silicon and magnesium are less important for WHIM observations as their observable transitions trace lower temperature gas (e.g., C IV, Si IV) or the abundance of these elements in the intergalactic medium are too low. It is important to note at this point, that the discussed relation between ionisation state/fraction and gas temperature explicitly assumes that the gas is in an ionisation *equilibrium*. This may not be generally the case in the WHIM, however,

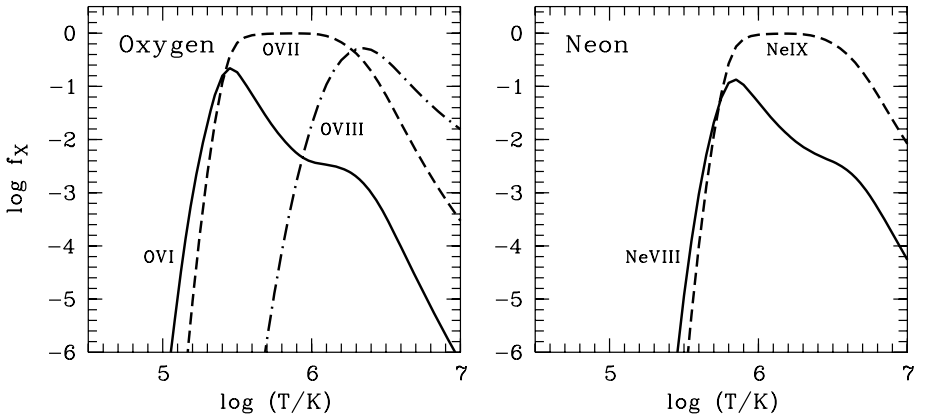


Fig. 3 CIE ion fractions of selected high ions of oxygen (O VI, O VII, O VIII; *left panel*) and neon (Ne VIII, Ne IX; *right panel*) in the WHIM temperature range $\log(T/\mathrm{K}) = 4.5-7.0$, based on calculations by Sutherland and Dopita (1993)

as the densities are generally very low. For instance, under particular non-equilibrium conditions the timescales for cooling, recombination, and ion/electron equilibration may differ significantly from each other (see for instance Bykov et al. 2008—Chap. 8, this issue). In such a case, the presence of high ions such as O VI and/or measured high-ion ratios would *not* serve as a reliable "thermometer" for the WHIM gas. In addition, WHIM filaments most likely neither are isothermal nor do they have a constant particle density. In fact, as WHIM simulations demonstrate, WHIM absorbers seem to represent a mix of cooler photoionised and hotter collisionally ionised gas with a substantial intrinsic density range. The absorption features from high ions arising in such a multi-phase medium therefore are generally difficult to interpret in terms of physical conditions and baryon budget.

In view of the high energies required to produce the high ions of oxygen and neon in combination with the spectral shape of the metagalactic background radiation (see Fig. 1), photoionisation of high metal ions in the WHIM is less important than for hydrogen. However, for O VI photoionisation is important at low redshifts in WHIM regions with very low densities or in systems located close to a strong local radiation source (e.g., in O VI systems associated with the background QSO). Note that at high redshift, most of the intervening O VI appears to be photoionised owing to the significantly higher intensity of the metagalactic background radiation in the early Universe (see Sect. 5.2).

2.2 WHIM Absorption Signatures in the UV and X-Ray Band

2.2.1 UV Absorption

As indicated in the previous subsection, five-times ionised oxygen (O VI) is by far the most important high ion to trace the WHIM at temperatures of $T \sim 3 \times 10^5$ K in the ultraviolet regime (assuming CIE, see above). Oxygen is a relatively abundant element and the two lithium-like $(1s^2 2s)\,^2S_{1/2} \rightarrow (1s^2 2p)\,^2P_{1/2,3/2}$ electronic transitions of O VI located in the FUV at 1031.9 and 1037.6 Å have large oscillator strengths ($f_{1031} = 0.133$, $f_{1037} = 0.066$). Next to O VI, Ne VIII traces WHIM gas near $T \sim 7 \times 10^5$ K (in collisional ionisation equilibrium) and thus is possibly suited to complement the O VI measurements of the WHIM in a higher temperature regime. The two available Ne VIII lines are located in the extreme

Table 1 Data on O and Ne high ions having observable absorption lines

Ion	[X/H][a]	Ionisation potential [eV]	Absorption lines [Å]	Band	CIE temperature[b] range [10^6 K]
O VI	−3.34	138	1031.926	FUV	0.2–0.5
			1037.617		
O VII	−3.34	739	21.602	X-ray	0.3–3.0
O VIII	−3.34	871	18.969	X-ray	1.0–10.0
Ne VIII	−4.16	239	770.409	EUV	0.5–1.3
			780.324		
Ne IX	−4.16	1196	13.447	X-ray	0.6–6.3

[a][X/H] is the log of the number density of element X relative to hydrogen for Solar abundances, taken here from Asplund et al. (2004)

[b]CIE models from Sutherland and Dopita (1993)

ultraviolet (EUV) at 770.4 Å ($f_{770} = 0.103$) and 780.3 Å ($f_{780} = 0.051$), allowing us to trace high-column density WHIM absorbers at redshifts $z > 0.18$ with current FUV satellites such as FUSE. However, as the cosmic abundance of Ne VIII is relatively low, Ne VIII is not expected to be a particularly sensitive tracer of the WHIM at the S/N levels achievable with current UV spectrographs. The same argument holds for the high ion Mg X, which has two transitions in the EUV at even lower wavelengths (λλ 609.8, 624.9 Å). So far, only O VI and in one case Ne VIII has been observed in the WHIM at low redshift (see Sect. 3.2). Note that WHIM absorption features by O VI (and Ne VIII) are mostly unsaturated and the line profiles are fully or nearly resolved by current UV instruments such as FUSE and STIS, which provide spectral resolutions of $R = \lambda/\Delta\lambda \approx 20{,}000$ and 45,000, respectively. Table 1 summarises physical parameters of O and Ne high ions and their observable transitions in the UV and X-ray bands.

Four-times ionised nitrogen (N V; I.P. is 98 eV) also is believed to trace predominantly collisionally ionised gas at temperatures near $T \sim 2 \times 10^5$ K, but its lower cosmic abundance together with its deficiency in low metallicity environments due to nucleosynthesis effects (e.g., Pettini et al. 2002) makes it very difficult to detect in the WHIM. Other available strong high-ion transitions in the UV from carbon (C IV λλ 1548.2, 1550.8 Å) and silicon (Si IV λλ 1393.8, 1402.8 Å) are believed to trace mainly photoionised gas at temperatures $T < 10^5$ K, but not the shock-heated warm-hot gas at higher temperatures.

Next to high-ion absorption from heavy elements, recent UV observations (Richter et al. 2004; Sembach et al. 2004; Lehner et al. 2007) have indicated that WHIM filaments can be detected in Ly α absorption of neutral hydrogen. Although the vast majority of the hydrogen in the WHIM is ionised (by collisional processes and UV radiation), a tiny fraction ($f_{\rm HI} < 10^{-5}$, typically) of neutral hydrogen is expected to be present. Depending on the total gas column density of a WHIM absorber and its temperature, weak H I Ly α absorption at column densities $12.5 \leq \log N({\rm H\,I}) \leq 14.0$ may arise from WHIM filaments and could be used to trace the ionised hydrogen component. The Ly α absorption from WHIM filaments is very broad due to thermal line broadening, resulting in large Doppler parameters of $b > 40$ km s^{-1}. Such lines are generally difficult to detect, as they are broad and shallow. High resolution, high S/N FUV spectra of QSOs with smooth background continua are required to successfully search for broad Ly α absorption in the low-redshift WHIM. STIS installed

on the HST is the only instrument that has provided such data, but due to the instrumental limitations of space-based observatories, the number of QSO spectra adequate for searching for WHIM broad Ly α absorption (in the following abbreviated as "BLA") is very limited.

The b values of the BLAs are assumed to be composed of a thermal component, $b_{\rm th}$, and a non-thermal component, $b_{\rm nt}$, in the way that

$$b = \sqrt{{b_{\rm th}}^2 + {b_{\rm nt}}^2}. \tag{7}$$

The non-thermal component may include processes like macroscopic turbulence, unresolved velocity-components, and others (see Richter et al. 2006a for a detailed discussion). The contribution of the thermal component to b depends on the gas temperature:

$$b_{\rm th} = \sqrt{\frac{2kT}{m}} \approx 0.13 \sqrt{\frac{T}{A}}\ {\rm km\,s^{-1}}, \tag{8}$$

where T is in K, k is the Boltzmann constant, m is the particle mass, and A is the atomic weight. For the shock-heated WHIM gas with $\log T \geq 5$ one thus expects $b_{\rm th} \geq 40$ km s^{-1}. The non-thermal broadening mechanisms are expected to contribute to some degree to the total b values in WHIM absorbers (see Richter et al. 2006a), so that the measured b value of a BLA provides only an upper limit for the temperature of the gas.

2.2.2 X-Ray Absorption

The highest ionisation phase of the WHIM will produce and absorb line radiation primarily in the He- and H-like ions of the low-Z elements (C, N, O, Ne), and possibly in the L-shell ions of Fe. In practice, much of the attention is focused on oxygen, because of its relatively high abundance, and because the strongest resonance lines in He- and H-like O are in a relatively 'clean' wavelength band. For reference, the Lyα transitions of C VI, N VII, O VIII, and Ne X occur at 33.7360, 24.7810, 18.9689, and 12.1339 Å, respectively (wavelengths of the $1s-2p_{1/2,3/2}$ doublet weighted with oscillator strength; Johnson and Soff 1985). The He-like ions C V, N VI, O VII, and Ne IX have their strongest transition, the $n = 1-2$ resonance line, at 40.2674, 28.7800, 21.6015, and 13.4473 Å (Drake 1988; see also Table 1). Data on the higher order series members can be found in Verner et al. (1996). As far as the Fe L shell ions are concerned, the most likely transition to show up would be the strongest line in Ne-like Fe XVII, $n = 2p-3d$ λ15.014 Å. In addition, all lower ionisation stages of C, N, O, and Ne (with the exception of neutral Ne of course) can also absorb by $n = 1-2$; the strongest of these transitions would be $1s-2p$ in O VI at 22.019 Å (Schmidt et al. 2004). Likewise, the lower ionisation stages of Fe could in principle produce $n = 2-3$ absorption.

The thermal widths of all these transitions will be very small, requiring resolving powers of order $R \sim 10\,000$ (C, N, O, Ne) for gas temperatures of order 10^6 K to be resolved; for Fe, the requirement is even higher, by a factor ~ 2. As we will see, for practical reasons, these requirements exceed the current capabilities of astrophysical X-ray spectroscopy by a large factor. Due to the small Doppler broadening (ignoring turbulent velocity fields for now), the lines will rapidly saturate. For He- and H-like O resonance line absorption, saturation sets in at an equivalent width of order 1 mÅ (Kaastra et al. 2008—Chap. 9, this issue) or column densities of order a few times 10^{14} ions cm^{-2}. The challenge, therefore, for X-ray spectroscopy presented by the IGM is to detect small equivalent width, near-saturation lines that are unresolved.

2.3 The Baryon Content of the WHIM as Measured by UV and X-Ray Absorbers

One important result from absorption line measurements of the WHIM in the UV is the observed number density of WHIM absorbers, usually expressed as $\mathrm{d}N/\mathrm{d}z$, the number of absorbers per unit redshift. For instance, from recent measurements with FUSE and HST/STIS one finds for O VI absorbers and Broad Ly α Absorbers at $z \approx 0$ values of $\mathrm{d}N/\mathrm{d}z(\mathrm{O\ VI}) \approx 20$ and $\mathrm{d}N/\mathrm{d}z(\mathrm{BLA}) \approx 30$ (see Sect. 3.2). Currently, the WHIM absorber density is only measurable in the UV, since in the X-ray band both the observed number of WHIM absorption lines and the available redshift path for WHIM observations is too small to derive statistically significant values of $\mathrm{d}N/\mathrm{d}z$(O VII) and $\mathrm{d}N/\mathrm{d}z$(O VIII).

A particularly interesting question now is, how the observed number density of high-ion lines or BLAs translates into an estimate of the cosmological baryon mass density of the WHIM, Ω_b(WHIM). To obtain such an estimate of the baryon content of the WHIM from UV and X-ray absorption measurements one has to consider two main steps. First, one needs to transform the observed column densities of the high ions (e.g., O VI, O VII, O VIII) into a total gas column density by modelling the ionisation conditions in the gas. In a second step, one then integrates over the total gas column densities of all observed WHIM absorbers along the given redshift path and from that derives Ω_b(WHIM) for a chosen cosmology. Throughout the paper we will assume a ΛCDM cosmology with $H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\Lambda = 0.7$, $\Omega_\mathrm{m} = 0.3$, and $\Omega_\mathrm{b} = 0.045$. For the first step the uncertainty lies in the estimate of the ionisation fraction of hydrogen of the WHIM. For this, it is usually assumed that the WHIM is in collisional ionisation equilibrium, but photoionisation and non-equilibrium conditions may play a significant role. In the case of using metal ions such as O VI the unknown oxygen abundance (O/H) of the gas introduces an additional uncertainty (see below) for the estimate of Ω_b(WHIM). For the second step, it is important to have a large enough sample of WHIM absorption lines and a sufficient total redshift path along *different* directions in order to handle statistical errors and the problem of cosmic variance. As mentioned earlier, these requirements currently are fulfilled only for the UV absorbers.

The cosmological mass density Ω_b of O VI absorbers (and, similarly, for other high ions) in terms of the current critical density ρ_c can be estimated by

$$\Omega_\mathrm{b}(\mathrm{O\ VI}) = \frac{\mu\, m_\mathrm{H}\, H_0}{\rho_\mathrm{c}\, c} \sum_{ij} \frac{N(\mathrm{O\ VI})_{ij}}{f_{\mathrm{O\ VI},ij}\,(\mathrm{O/H})_{ij}\,\Delta X_j}. \tag{9}$$

In this equation, $\mu = 1.3$ is the mean molecular weight, $m_\mathrm{H} = 1.673 \times 10^{-27}$ kg is the mass per hydrogen atom, H_0 is the adopted local Hubble constant, and $\rho_\mathrm{c} = 3{H_0}^2/8\pi G$ is the current critical density. The index i denotes an individual high-ion absorption system along a line of sight j. Each measured high-ion absorption system i is characterised by its measured ion column density (e.g., $N(\mathrm{O\ VI})_{ij}$), the ionisation fraction of the measured ion (e.g., $f_{\mathrm{O\ VI},ij}$), and the local abundance of the element measured compared to hydrogen (e.g., the local oxygen-to-hydrogen ratio, by number). Each line of sight j has a characteristic redshift range Δz in which high-ion absorption may be detected. The corresponding comoving path length ΔX available for the detection of WHIM high-ion absorbers then is given by:

$$\Delta X_j = (1+z)^2\,[\Omega_\Lambda + \Omega_\mathrm{m}(1+z)^3]^{-0.5}\,\Delta z_j. \tag{10}$$

In analogy, we can write for the cosmological mass density of the BLAs:

$$\Omega_\mathrm{b}(\mathrm{BLA}) = \frac{\mu\, m_\mathrm{H}\, H_0}{\rho_\mathrm{c}\, c} \sum_{ij} \frac{N(\mathrm{H\ I})_{ij}}{f_{\mathrm{H\ I},ij}\,\Delta X_j}. \tag{11}$$

As can be easily seen, the advantage of using BLAs for deriving the WHIM mass density is that the metallicity of the gas is unimportant for the determination of Ω_b. The disadvantage is, however, that the ionisation corrections are very large and uncertain, since they are determined indirectly from the BLA line widths (see Sect. 2.2.1).

3 UV Measurements of the WHIM

3.1 Past and Present UV Instruments

The first and second generations of space based UV spectrographs such as *Copernicus* and the *International Ultraviolet Explorer* (IUE) did not have sufficient sensitivity to systematically study intervening absorption in the intergalactic medium along a large number of sightlines. The early low- and intermediate resolution spectrographs installed on the *Hubble Space Telescope* (HST), namely the *Faint Object Spectrograph* (FOS) and the *Goddard High Resolution Spectrograph* (GHRS), were used to study the properties of the local Ly α forest and intervening metal-line systems (e.g., Stocke et al. 1995; Shull et al. 1998). While intervening O VI absorption has been detected with these instruments (e.g., Tripp et al. 1998), the concept of a warm-hot intergalactic gas phase was not really established at that time. With the implementation of the high-resolution capabilities of the *Space Telescope Imaging Spectrograph* (STIS) installed on HST the first systematic analyses of WHIM O VI absorbers as significant low-redshift baryon reservoirs came out in 2000 (see Tripp et al. 2000), thus relatively soon after the importance of a shock-heated intergalactic gas phase was realised in cosmological simulations for the first time (e.g., Cen and Ostriker 1999; Davé et al. 2001). The STIS echelle spectrograph together with the E140M grating provides a high spectral-resolution of $R \approx 45\,000$, corresponding to a velocity resolution of ~ 7 km s^{-1} in the STIS E140M wavelength band between 1150 and 1730 Å (e.g., Kimble et al. 1998; Woodgate 1998). An example for a STIS quasar spectrum with intervening hydrogen and metal-line absorption is shown in Fig. 4. Note that at the spectral resolution of the STIS E140M grating all intergalactic absorption lines (i.e., hydrogen and metal lines) are fully resolved. In 1999, the *Far Ultraviolet Spectroscopic Explorer* (FUSE) became available, covering the wavelength range between 912 and 1187 Å. Equipped with a Rowland-type spectrograph providing a medium spectral resolution of $R \approx 20\,000$ (FWHM $\sim$ 20 km s^{-1}) FUSE is able to observe extragalactic UV background sources brighter than $V = 16.5$ mag with acceptable integration time and signal-to-noise (S/N) ratios (for a description of FUSE see Moos et al. 2000; Sahnow et al. 2000). With this resolution, FUSE is able to resolve the broader intergalactic absorption from the H I Lyman series, while most of the narrow metal-line absorbers remain just unresolved. This is not a problem for O VI WHIM studies with FUSE, however, since the spectral resolution is very close to the actual line widths and the O VI absorption usually is not saturated. FUSE complements the STIS instruments at lower wavelengths down to the Lyman limit and consequently combined FUSE and STIS spectra of ~ 15 low redshift QSOs and AGN have been used to study the low-redshift WHIM via intervening O VI and BLA absorption (see Tripp et al. 2007 and references therein). Unfortunately, since 2006/2007 both STIS and FUSE are out of commission due to technical problems.

Fresh spectroscopic UV data from WHIM absorption line studies will become available once the *Cosmic Origins Spectrograph* (COS) will be installed on HST during the next HST

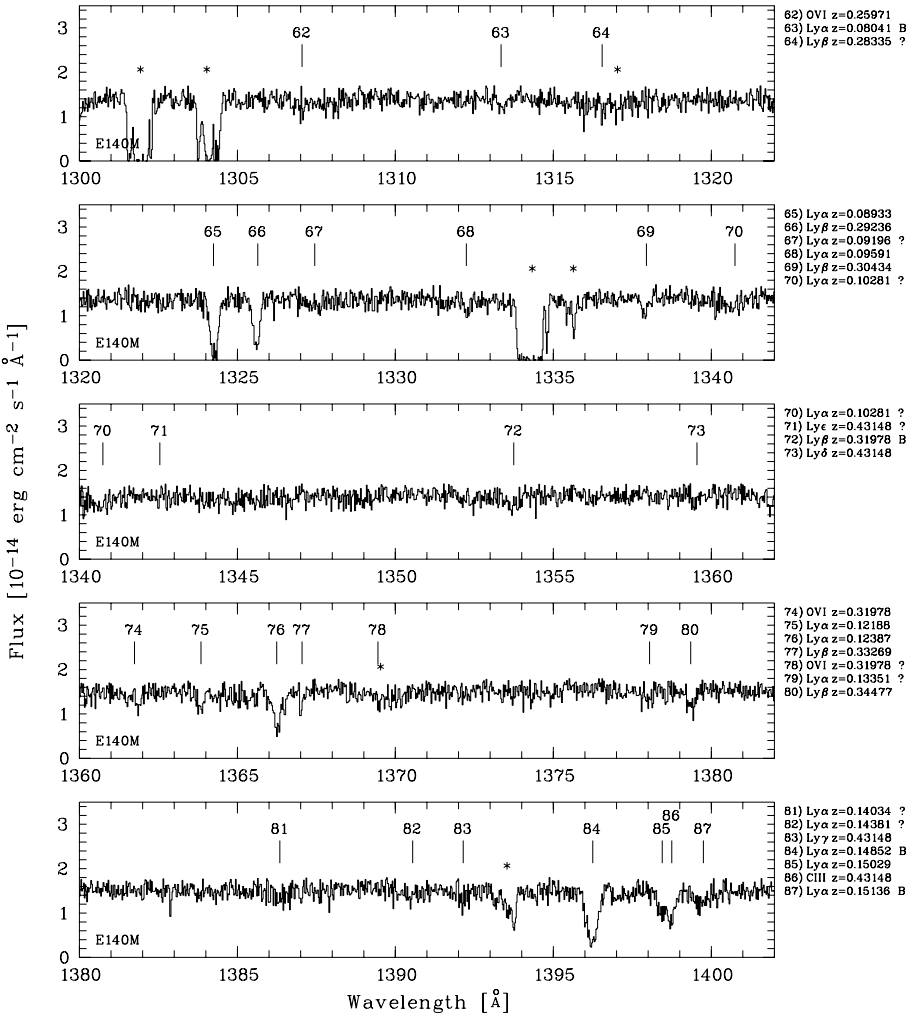


Fig. 4 STIS spectrum of the quasar PG 1259+593 in the wavelength range between 1300 and 1400 Å. Next to absorption from the local Ly α forest and gas in the Milky Way there are several absorption features that most likely are related to highly ionised gas in the WHIM. Absorption from five-times ionised oxygen (O VI) is observed at $z = 0.25971$ and $z = 0.31978$. Broad H I Ly α and Ly β absorption is detected at $z = 0.08041$, 0.09196, 0.10281, 0.13351, 0.14034, 0.14381, 0.14852, 0.15136, and $z = 0.31978$. From Richter et al. (2004)

service mission (SM-4), which currently is scheduled for late 2008. COS will observe in the UV wavelength band between 1150 and 3000 Å at medium resolution ($R \approx 20\,000$). COS has been designed with maximum effective area as the primary constraint: it provides more than an order of magnitude gain in sensitivity over previous HST instruments. Due to its very high sensitivity, COS thus will be able to observe *hundreds* of low- and intermediate redshift QSOs and AGN and thus will deliver an enormous data archive to study the properties of WHIM UV absorption lines systems in great detail (see also Paerels et al. 2008—Chap. 19, this issue).

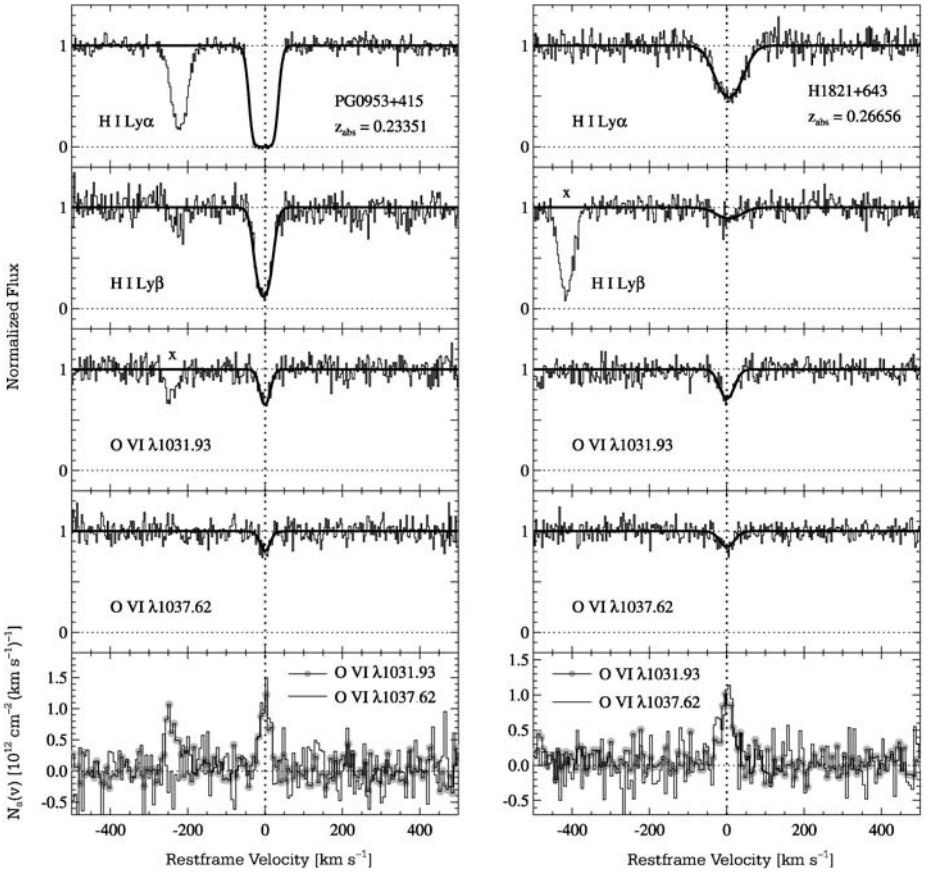


Fig. 5 Examples for H I and O VI absorption in two absorption systems at $z = 0.23351$ and $z = 0.6656$ towards PG 0953+415 and H 1821+643, respectively, plotted on a rest frame velocity scale (observed with STIS). Adapted from Tripp et al. (2007)

3.2 Intervening WHIM Absorbers at Low Redshift

We start with the O VI absorbers which are believed to trace the low-temperature tail of the WHIM at $T < 5 \times 10^5$ K. Up to now, more than 50 detections of intervening O VI absorbers at $z < 0.5$ have been reported in the literature (e.g., Tripp et al. 2000; Oegerle et al. 2000; Chen and Prochaska 2000; Savage et al. 2002; Richter et al. 2004; Sembach et al. 2004; Savage et al. 2005; Danforth and Shull 2005; Tripp et al. 2007). All of these detections are based on FUSE and STIS data. Figure 5 shows two examples for intervening O VI absorption at $z = 0.23351$ and $z = 0.26656$ in the direction of PG 0953+415 and H 1821+643, as observed with STIS. The most recent compilation of low-redshift intervening O VI absorbers is that of Tripp et al. (2007), who have analysed 16 sightlines toward low-redshift QSOs observed with STIS and FUSE along a total redshift path of $\Delta z \approx 3$. They find a total of 53 intervening O VI absorbers (i.e., they are not within 5000 km s^{-1} of z_{QSO}) comprised of 78 individual absorption components.[1] The measurements imply a number density of O VI ab-

[1] I.e., some of the O VI systems have velocity sub-structure.

sorbing systems per unit redshift of $dN_{O\,VI}/dz \approx 18 \pm 3$ for equivalent widths $W_\lambda \geq 30$ mÅ. The corresponding number density of O VI absorption *components* is $dN_{O\,VI}/dz \approx 25 \pm 3$. These values are slightly higher than what is found by earlier analyses of smaller O VI samples (Danforth and Shull 2005), but lie within the cited 2σ error ranges. The discrepancy between the measured O VI number densities probably is due to the different approaches of estimating the redshift path Δz along which the O VI absorption takes place. If one assumes that the gas is in a collisional ionisation equilibrium, i.e., that ~ 20 percent of the oxygen is present in the form of O VI ($f_{O\,VI} \leq 0.2$), and further assumes that the mean oxygen abundance is 0.1 Solar, the measured number density of O VI absorbers corresponds to a cosmological mass density of $\Omega_b(O\,VI) \approx 0.0020$–$0.0030\ {h_{70}}^{-1}$. These values imply that intervening O VI absorbers trace $\sim 5-7$ percent of the total baryon mass in the local Universe. For the interpretation of $\Omega_b(O\,VI)$ it has to be noted that O VI absorption traces collisionally ionised gas at temperatures around 3×10^5 K (and also low-density, photoionised gas at lower temperatures), but not the million-degree gas phase which probably contains the majority of the baryons in the WHIM.

The recent analysis of Tripp et al. (2007) indicates, however, that this rather simple conversion from measured O VI column densities to $\Omega_b(O\,VI)$ may not be justified in general, as the CIE assumption possibly breaks down for a considerable fraction of the O VI systems. From the measured line widths of the H I Ly α absorption that is associated with the O VI Tripp et al. conclude that ~ 40 percent of their O VI systems belong to cooler, photoionised gas with $T < 10^5$ K, possibly not at all associated with shock-heated warm-hot gas. In addition, about half of the intervening O VI absorbers arise in rather complex, multi-phase systems that can accommodate hot gas at relatively low metallicity. It thus appears that—without having additional information about the physical conditions in each O VI absorber—the estimate of the baryon budget in intervening O VI systems is afflicted with rather large systematic uncertainties.

In high-column density O VI systems at redshifts $z > 0.18$, such desired additional information may be provided by the presence or absence of Ne VIII (see Sect. 2.2.1), which in CIE traces gas at $T \sim 7 \times 10^5$ K. Toward the quasar PG 1259+593 Richter et al. (2004) have reported a tentative detection of Ne VIII absorption at $\sim 2\sigma$ significance in an O VI absorber at $z \approx 0.25$. The first secure detection of intervening Ne VIII absorption (at $\sim 4\sigma$ significance) was presented by Savage et al. (2005) in a multi-phase O VI absorption system at $z \approx 0.21$ in the direction of the quasar HE 0226−4110. The latter authors show that in this particular absorber the high-ion ratio Ne VIII/O VI $= 0.33$ is in agreement with gas in CIE at temperature of $T \sim 5 \times 10^5$ K. With future high S/N absorption line data of low-redshift QSOs (as will be provided by COS) it is expected that the number of detections of WHIM Ne VIII absorbers will increase substantially, so that an important new diagnostic will become available for the analysis of high-ion absorbers.

One other key aspect in understanding the distribution and nature of intervening O VI systems concerns their relation to the large-scale distribution of galaxies. Combining FUSE data of 37 O VI absorbers with a database of more than a million galaxy positions and redshifts, Stocke et al. (2006) find that all of these O VI systems lie within $800\,{h_{70}}^{-1}$ kpc of the nearest galaxy. These results suggest that O VI systems preferentially arise in the immediate circumgalactic environment and extended halos of galaxies, where the metallicity of the gas is expected to be relatively high compared to regions far away from galactic structures. Some very local analogs of intervening O VI systems thus may be the O VI high-velocity clouds in the Local Group that are discussed in the next subsection. Due to apparent strong connection between intervening O VI systems and galactic structures and a resulting galaxy/metallicity bias problem it is of great interest to consider other tracers of warm-hot gas, which are

independent of the metallicity of the gas. The broad hydrogen Lyα absorbers—as will be discussed in the following—therefore represent an important alternative for studying the WHIM at low redshift.

As described in Sect. 2.2.1, BLAs represent H I Lyα absorbers with large Doppler parameters $b > 40$ km s^{-1}. If thermal line broadening dominates the width of the absorption, these systems trace the WHIM at temperatures between 10^5 and 10^6 K, typically (note that for most systems with $T > 10^6$ K BLAs are both too broad and too shallow to be unambiguously identified with the limitations of current UV spectrographs). The existence of H I Lyα absorbers with relatively large line widths has been occasionally reported in earlier absorption-line studies of the local intergalactic medium (e.g., Tripp et al. 2001; Bowen et al. 2002). Motivated by the rather frequent occurrence of broad absorbers along QSO sightlines with relatively large redshift paths, the first systematic analyses of BLAs in STIS low-z data were carried out by Richter et al. (2004) and Sembach et al. (2004). Richter et al. (2006a) have inspected four sightlines observed with STIS towards the quasars PG 1259+593 ($z_{\rm em} = 0.478$), PG 1116+215 ($z_{\rm em} = 0.176$), H 1821+643 ($z_{\rm em} = 0.297$), and PG 0953+415 ($z_{\rm em} = 0.239$) for the presence of BLAs and they identified a number of good candidates. Their study implies a BLA number density per unit redshift of $dN_{\rm BLA}/dz \approx 22 - 53$ for Doppler parameters $b \geq 40$ km s^{-1} and above a sensitivity limit of $\log(N({\rm cm}^{-2})/b({\rm km\,s}^{-1})) \geq 11.3$. The large range for $dN_{\rm BLA}/dz$ partly is due to the uncertainty about defining reliable selection criteria for separating spurious cases from good broad Lyα candidates (see discussions in Richter et al. 2004, 2006a and Sembach et al. 2004). Transforming the number density $dN_{\rm BLA}/dz$ into a cosmological baryonic mass density, Richter et al. (2006a) obtains $\Omega_{\rm b}({\rm BLA}) \geq 0.0027\,{h_{70}}^{-1}$. This limit is about 6 percent of the total baryonic mass density in the Universe expected from the current cosmological models (see above), and is comparable with the value derived for the intervening O VI absorbers (see above). Examples for several BLAs in the STIS spectrum of the quasar H 1821+643 are shown in Fig. 6.

More recently, Lehner et al. (2007) have analysed BLAs in low-redshift STIS spectra along seven sightlines. They find a BLA number density of $dN_{\rm BLA}/dz = 30 \pm 4$ for $b = 40{-}150$ km s^{-1} and $\log N({\rm H\,I}) > 13.2$ for the redshift range $z = 0{-}0.4$. They conclude that BLAs host at least 20 percent of the baryons in the local Universe, while the photoionised Lyα forest, which produces a large number of narrow Lyα absorbers (NLAs), contributes with ~ 30 percent to the total baryon budget. In addition, Prause et al. (2007) have investigated the properties of BLAs at intermediate redshifts ($z = 0.9{-}1.9$) along five other quasars using STIS high- and intermediate-resolution data. They find a number density of reliably detected BLA candidates of $dN_{\rm BLA}/dz \approx 14$ and obtain a lower limit of the contribution of BLAs to the total baryon budget of ~ 2 percent in this redshift range. The frequency and baryon content of BLAs at intermediate redshifts obviously is lower than at $z = 0$, indicating that at intermediate redshifts shock-heating of the intergalactic gas from the infall in large-scale filaments is not yet very efficient. This is in line with the predictions from cosmological simulations.

3.3 The Milky Way Halo and Local Group Gas

One primary goal of the FUSE mission was to constrain the distribution and kinematics of hot gas in the thick disk and lower halo of the Milky Way by studying the properties of Galactic O VI absorption systems at radial velocities $|v_{\rm LSR}| \leq 100$ km s^{-1} (Savage et al. 2000, 2003; Wakker et al. 2003). However, as the FUSE data unveil, O VI absorption associated with Milky Way gas is observed not only at low velocities but also at $|v_{\rm LSR}| > 100$

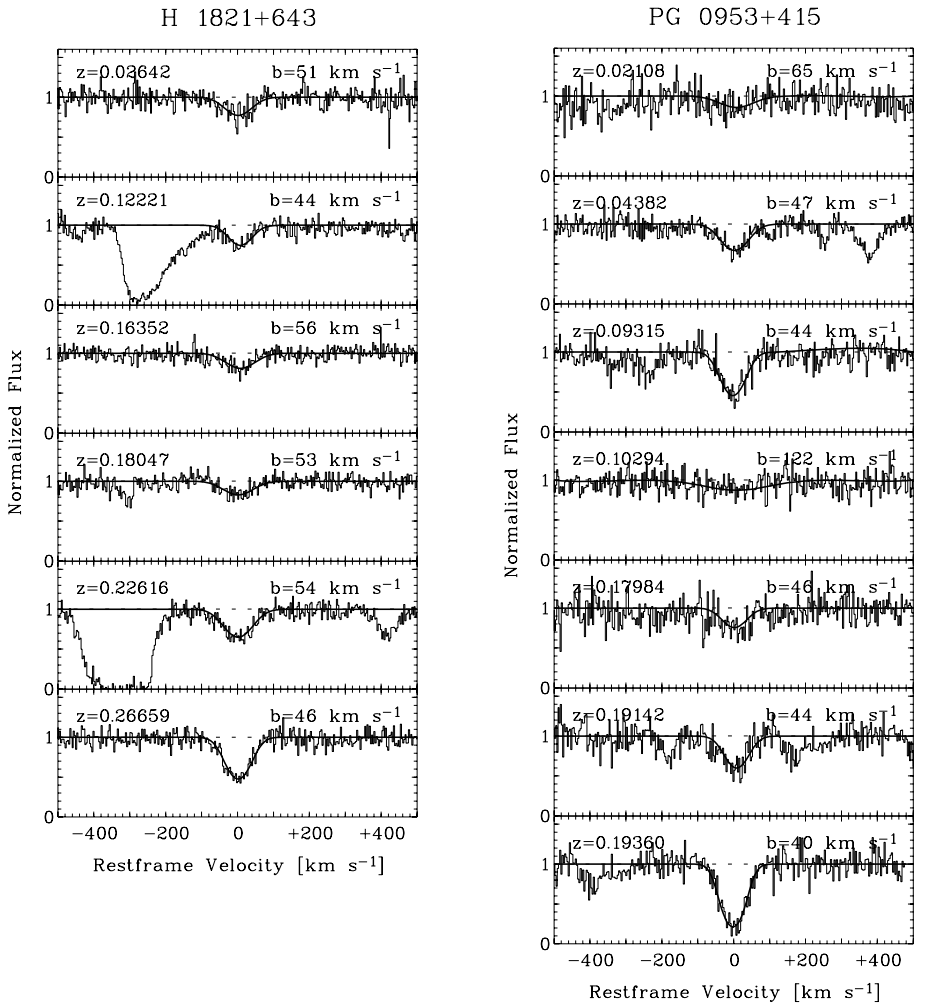


Fig. 6 Broad Lyman α absorbers towards the quasars H 1821+643 and PG 0953+415 (STIS observations), plotted on a rest frame velocity scale. If thermal line broadening dominates the width of the absorption, these systems trace the WHIM at temperatures between 10^5 and 10^6 K. From Richter et al. (2006a)

km s^{-1} (Sembach et al. 2003). The topic of cool and hot gas in the halo of the Milky Way recently has been reviewed by Richter (2006c). These detections imply that next to the Milky Way's hot "atmosphere" (i.e., the Galactic Corona) individual pockets of hot gas exist that move with high velocities through the circumgalactic environment of the Milky Way. Such high-velocity O VI absorbers may contain a substantial fraction of the baryonic matter in the Local Group in the form of warm-hot gas and thus—as discussed in the previous subsection—possibly represent the local counterparts of some of the intervening O VI absorbers observed towards low-redshift QSOs.

From their FUSE survey of high-velocity O VI absorption Sembach et al. (2003) find that probably more than 60 percent of the sky at high velocities is covered by ionised hydrogen (associated with the O VI absorbing gas) above a column density level of $\log N(\mathrm{H\,II}) = 18$,

assuming a metallicity of the gas of 0.2 Solar. Some of the high-velocity O VI detected with FUSE appears to be associated with known high-velocity H I 21 cm structures (e.g., the high-velocity clouds complex A, complex C, the Magellanic Stream, and the Outer Arm). Other high-velocity O VI features, however, have no counterparts in H I 21 cm emission. The high radial velocities for most of these O VI absorbers are incompatible with those expected for the hot coronal gas (even if the coronal gas motion is decoupled from the underlying rotating disk). A transformation from the Local Standard of Rest to the Galactic Standard of Rest and the Local Group Standard of Rest velocity reference frames reduces the dispersion around the mean of the high-velocity O VI centroids (Sembach et al. 2003; Nicastro et al. 2003). This can be interpreted as evidence that *some* of the O VI high-velocity absorbers are intergalactic clouds in the Local Group rather than clouds directly associated with the Milky Way. However, it is extremely difficult to discriminate between a Local Group explanation and a distant Galactic explanation for these absorbers. The presence of intergalactic O VI absorbing gas in the Local Group is in line with theoretical models that predict that there should be a large reservoir of hot gas left over from the formation of the Local Group (see, e.g., Cen and Ostriker 1999).

It is unlikely that the high-velocity O VI is produced by photoionisation. Probably, the gas is collisionally ionised at temperatures of several 10^5 K. The O VI then may be produced in the turbulent interface regions between very hot ($T > 10^6$ K) gas in an extended Galactic Corona and the cooler gas clouds that are moving through this hot medium (see Sembach et al. 2003). Evidence for the existence of such interfaces also comes from the comparison of absorption lines from neutral and weakly ionised species with absorption from high ions like O VI (Fox et al. 2004).

4 X-Ray Measurements of the WHIM

4.1 Past and Present X-Ray Instruments

With the advent of the *Chandra* and *XMM-Newton* observatories, high resolution X-ray spectroscopy of a wide variety of cosmic sources became feasible for the first time. Among the possible results most eagerly speculated upon was the detection of intergalactic absorption lines from highly ionised metals in the continuum spectra of bright extragalactic sources. After all, one of the most striking results from the *Einstein* observatory had been the detection of a very significant broad absorption feature at ~ 600 eV in the spectrum of PKS 2155−304 with the Objective Grating Spectrometer (Canizares and Kruper 1984). Ironically, if interpreted as intergalactic H-like O Lyα absorption, redshifted and broadened by the expansion of the Universe, the strength of the feature implied the presence of a highly ionised IGM of near-critical density, a possibility that has of course definitively been discounted since then.

The High Energy Transmission Grating Spectrometer (HETGS; Canizares et al. 2005) and the Low Energy Transmission Grating Spectrometer (LETGS; Brinkman et al. 2000) on *Chandra*, and the Reflection Grating Spectrometer (RGS) on *XMM-Newton* (den Herder et al. 2001) were the first instruments to provide sensitivity to weak interstellar and intergalactic X-ray absorption lines. The 'traditional' ionisation detectors (proportional counters, CCD's) do not have sufficient energy resolution for this application. But the angular resolution provided by an X-ray telescope can be used to produce a high resolution spectrum, by the use of diffracting elements. Laboratory X-ray spectroscopy is typically performed with crystal diffraction spectrometers, and the use of crystal spectrometers for general use

in astrophysics was pioneered on the *Einstein* observatory (e.g., Canizares et al. 1979). The Focal Plane Crystal Spectrometer indeed detected the first ever narrow X-ray absorption line in a cosmic source, the $1s-2p$ absorption by neutral oxygen in the interstellar medium towards the Crab (Canizares and Kruper 1984). Previous grating spectrometers (the Objective Grating Spectrometer on *Einstein* and the two Transmission Grating Spectrometers on *EXOSAT*) had only limited resolution and sensitivity. But there is no fundamental limit to the resolution of a diffraction grating spectrometer, and the high angular resolution of the *Chandra* telescope has allowed for high resolution spectroscopy using transmission gratings. The focusing optics on *XMM-Newton* have more modest angular resolution, but they are used with a fixed array of grazing incidence reflection gratings, which produce very large dispersion angles and thus high spectral resolution. The HETGS provides a spectral resolution of $\Delta\lambda = 0.0125$ Å over the $\approx 1.5-15$ Å band with the high line density grating, and $\Delta\lambda = 0.025$ Å over the $2-20$ Å band with the medium line density grating. The LETGS has $\Delta\lambda = 0.05$ Å over the $\approx 2-170$ Å band, while the RGS has $\Delta\lambda = 0.06$ Å over the $5-38$ Å band. These numbers translate to resolving powers of $R = 400-1500$ in the O K band, and with a sufficiently bright continuum source, one should be able to detect equivalent widths of order $0.1-0.5$ eV ($5-20$ mÅ), or below in spectra with very high signal-to-noise. The predictions for H- and He-like O resonance absorption line strengths are generally smaller than these thresholds, but not grossly so, and so a search for intergalactic O was initiated early on. Given that the current spectrometers are not expected to resolve the absorption lines, the only freedom we have to increase the sensitivity of the search is to increase the signal-to-noise ratio in the continuum, and it becomes crucial to find suitable, very bright sources, at redshifts that are large enough that there is a reasonable a priori probability of finding a filament with detectable line absorption.

As we will see when we discuss the results of the observational searches for X-ray absorption lines, the problem is made considerably more difficult by the very sparseness of the expected absorption signature. Frequently, when absorption features of marginal statistical significance are detected in astrophysical data, plausibility is greatly enhanced by simple, unique spectroscopic arguments. For instance, for all plausible parameter configurations, an absorbing gas cloud detected in N V should also produce detectable C IV absorption; or, both members of a doublet should appear in the correct strength ratio if unsaturated. In the early stages of Lyα forest astrophysics, it was arguments of this type, rather than the crossing of formal statistical detection thresholds alone, that guided the field (e.g., Lynds 1971). But for the 'X-ray Forest' absorption, we expect a very different situation. The detailed simulations confirm what simple analytical arguments had suggested: in most cases, Intergalactic absorption systems that are in principle detectable with current or planned X-ray instrumentation will show just a single absorption line, usually the O VII $n = 1-2$ resonance line, at an unknown redshift. When assessing the possibility that a given apparent absorption feature is 'real', one has to allow for the number of independent trial redshifts (very roughly given by the width of the wavelength band surveyed, divided by the nominal spectral resolution of the spectrometer), and with a wide band and a high spectral resolution, this tends to dramatically reduce the significance of even fairly impressive apparent absorption dips. For instance, an apparent absorption line detected at a formal '3σ significance' (or $p = 0.0015$ a priori probability for a negative deviation this large to occur due to statistical fluctuation) with *Chandra* LETGS in the $21.6-23.0$ Å band pales to '$1.7\,\sigma$' if we assume it is the O VII resonance line in the redshift range $z = 0-0.065$; with $N \approx 30$ independent trials, the chances of not seeing a 3σ excursion are $(1-p)^N = 0.956$, or: one will see such a feature one in twenty times if one tries this experiment (we are assuming a Gaussian distribution of fluctuations here). If we allow for confusion with O VIII Lyα at higher redshift, or even other transitions, the significance is even further reduced. And the larger the number of sources surveyed, the larger

the probability of false alarm. Clearly, more reliable statistics on intervening X-ray absorbers and detections at higher significance are desired, but the required high-quality data will not be available until the next-generation X-ray facilities such as *XEUS* and *Constellation X* are installed (see Paerels et al. 2008—Chap. 19, this issue).

Nevertheless, even with these odds, the above discussed high-ion measurements are important observations to do with the currently available instruments *Chandra* and *XMM-Newton*. Given the predicted strengths of the absorption lines (e.g., Chen et al. 2003; see also Sect. 5.1), attention has naturally focused on a handful of very bright BL Lac- and similar sources. Below, we discuss the results of the searches. Note that the subject has recently also been reviewed by Bregman (2007).

4.2 Intervening WHIM Absorbers at Low Redshift

Early observations of a sample of sources (e.g., S5 0836+710, PKS 2149−306; Fang et al. 2001; PKS 2155−304, Fang et al. 2002) produced no convincing detections. The next attempt at detecting redshifted X-ray O absorption lines was performed by Mathur et al. (2003) with a dedicated deep observation (470 ks) with the *Chandra* LETGS of the quasar H 1821+643, which has several confirmed intervening O VI absorbers. No significant X-ray absorption lines were found at the redshifts of the O VI systems, but this was not really surprising in view of the modest signal to noise ratio in the X-ray continuum. Since it requires very bright continua to detect the weak absorption, it is also not surprising that the number of suitable extragalactic sources is severely limited. Nicastro and his colleagues then embarked on a campaign to observe Mrk 421 during its periodic X-ray outbursts, when its X-ray flux rises by an order of magnitude (e.g., Nicastro 2005). The net result of this has been the accumulation of a very deep spectrum with the *Chandra* LETGS, with a total of more than 7 million continuum counts, in about 1000 resolution elements. Nicastro et al. (2005) have claimed evidence for the detection of two intervening absorption systems in these data, at $z = 0.011$ and $z = 0.027$. But the spectrum of the same source observed with the *XMM-Newton* RGS does not show these absorption lines (Rasmussen et al. 2007), despite higher signal-to-noise and comparable spectral resolution (Mrk 421 is observed by *XMM-Newton* for calibration purposes, and by late 2006, more than 1 Ms exposure had been accumulated). Kaastra et al. (2006) have reanalysed the *Chandra* LETGS data, and find no significant absorption. Other sources, less bright but with larger redshifts, have been observed (see for instance Steenbrugge et al. (2006) for observations of 1ES 1028+511 at $z = 0.361$), but to date no convincing evidence for intervening absorption has materialised.

Observations have been conducted to try and detect the absorption by intergalactic gas presumably associated with known locations of cosmic overdensity, centred on massive clusters. Fujimoto et al. (2004) attempted to detect absorption in the quasar LBQS 1228+1116, located behind the Virgo cluster. An *XMM-Newton* RGS spectrum revealed a marginal feature at the (Virgo) redshifted position of O VIII Lyα, but only at the $\sim 95\%$ confidence level. Likewise, Takei et al. (2007) took advantage of the location of X Comae behind the Coma cluster to try and detect absorption from Coma or its surroundings, but no convincing, strong absorption lines were detected in a deep observation with *XMM-Newton* RGS. The parallel CCD imaging data obtained with EPIC show weak evidence for Ne IX $n = 1-2$ line emission at the redshift of Coma, which, if real, would most likely be associated with WHIM gas around the cluster, seen in projection (the cluster virial temperature is too high for Ne IX). In practice, the absence of very bright point sources behind clusters, which makes absorption studies difficult, and the bright foregrounds in emission, will probably make this approach to detecting and characterising the WHIM not much easier than the random line-of-sight searches.

The conclusion from the search for intergalactic X-ray absorption is that there is no convincing, clear detection for intervening absorption. This is, in retrospect, not that surprising, given the sensitivity of the current X-ray spectrometers, the abundance of suitably bright and sufficiently distant continuum sources, and the predicted properties of the WHIM.

4.3 The Milky Way Halo and Local Group Gas

The first positive result of the analysis of bright continuum spectra was the detection of O VII and O VIII $n = 1-2$ resonance line absorption at redshift zero. Nicastro et al. (2002) first identified the resonance lines in the *Chandra* LETGS spectrum of PKS 2155−304 (O VII $n = 1-2$, $n = 1-3$, O VIII Lyα, Ne IX $n = 1-2$). Fang et al. (2003) found absorption in the LETGS spectrum of 3C273. Rasmussen et al. (2003) detected resonance absorption in the *XMM-Newton* RGS spectra of 3C 273, Mrk 421, and PKS 2155−304. Since then, at least O VII $n = 1-2$ has been detected in effectively all sufficiently bright continuum sources, both with *Chandra* and *XMM-Newton*; a recent compilation appears in Fang et al. (2006). Portions of a deeper spectrum that shows the zero redshift absorption are shown in Fig. 7.

Nicastro et al. (2002) initially interpreted the absorption as arising in an extended intergalactic filament. The argument that drives this interpretation is based on the assumption that O VI, O VII, and O VIII are all located in a single phase of the absorbing gas. The simultaneous appearance of finite amounts of O VI and O VIII only occurs in photoionised gas, not in

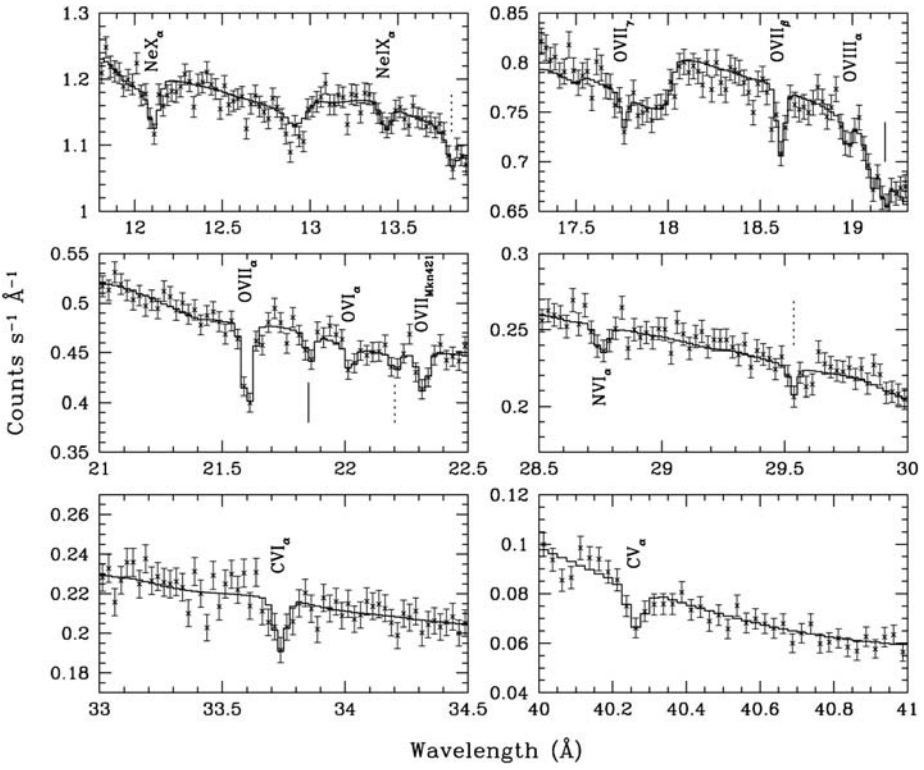


Fig. 7 *Chandra* LETGS spectrum of Mrk 421. Crosses are the data, the *solid line* is a model. The labels identify $z \approx 0$ absorption lines in Ne, O, and C. The *vertical tick marks* indicate the locations of possible intergalactic absorption lines. From Williams et al. (2005)

gas in collisional ionisation equilibrium, and this requires that the gas has very low density (the photoionisation is produced by the local X-ray background radiation field; for a measured ionisation parameter, the known intensity of the ionising field fixes the gas density). The measured ionisation balance then implies a length scale on the order of $l \sim 10\ (Z_{0.1})^{-1}$ Mpc, where $Z_{0.1}$ is the metallicity in units 0.1 Solar. This is a very large length, and even for 0.3 $Z_{\odot}$ metallicity, the structure still would not fit in the Local Group (and it is unlikely to have this high a metallicity if it were larger than the Local Group). In fact, the absorption lines should have been marginally resolved in this case, if the structure expands with a fair fraction of the expansion of the Universe.

Rasmussen et al. (2003) constrained the properties of the absorbing gas by dropping the O VI, and by taking into account the intensity of the diffuse O VII and O VIII line emission as measured by the Wisconsin/Goddard rocket-borne Quantum X-ray Calorimeter (XQC) experiment (McCammon et al. 2002). The cooling timescale of O VI-bearing gas is much smaller than that of gas with the higher ionisation stages of O, and this justifies the assumption that O VI is located in a different, transient phase of the gas. With only O VII and O VIII, the medium can be denser and more compact, and be in collisional ionisation equilibrium. Treating the measured O VII emission line intensity as an upper limit to the emission from a uniform medium, and constraining the ionisation balance from the measured ratio of O VII and O VIII column densities in the lines of sight to Mrk 421 and PKS 2155−304, Rasmussen et al. derived an upper limit on the density of the medium of $n_{\rm e} \lesssim 2 \times 10^{-4}$ cm^{-3} and a length scale $l \gtrsim 100$ kpc.

Bregman (2007) favours a different solution, with a lower electron temperature and hence a higher O VII ion fraction. If one assumes Solar abundance and sets the O VII fraction to 0.5, the characteristic density becomes $n_e \sim 10^{-3}$ cm^{-3}, and the length scale ($l \sim 20$ kpc) suggests a hot Galactic halo, rather than a Local Group intragroup medium.

Arguments for both type of solution (a small compact halo and a more tenuous Local Group medium) can be given. The most direct of these is a measurement of the O VII line absorption towards the LMC by Wang et al. (2005) in the spectrum of the X-ray binary LMC X-3, which indicates that a major fraction of the O VII column in that direction is in fact in front of the LMC. Bregman (2007) points out that the distribution of column densities of highly ionised O on the sky is not strongly correlated with the likely projected mass distribution of the Local Group, and that the measured velocity centroid of the absorption lines appears characteristic of Milky Way gas, rather than Local Group gas. On the other hand, a direct measurement of the Doppler broadening of the O VII gas, from the curve of growth of the $n = 1-2$ and $n = 1-3$ absorption lines in the spectrum of Mrk 421 and PKS 2155−304 (Williams et al. 2005, 2007), indicates an ion temperature of $T_{\rm i} \approx 10^{6.0-6.3}$ K (Mrk 421) and $T_i \approx 10^{6.2-6.4}$ K (PKS 2155−304), and these values favour the low-density, Local Group solution. Regardless, the prospect of directly observing hot gas expelled from the Galactic disk, or measuring the virial temperature of the Galaxy and/or the Local Group is exciting enough to warrant further attention to redshift zero absorption and emission.

Finally, the spectrum of Mrk 421 shows the expected $z = 0$ innershell O VI absorption, at 22.019 Å, both with *Chandra* and with *XMM-Newton*. There has been some confusion regarding an apparent discrepancy between the O VI column densities derived from the FUV and from the X-ray absorption lines, in the sense that the X-ray column appeared to be significantly larger than the FUV column (Williams et al. 2005). Proposed physical explanations for this effect involve a depletion of the lower level of the FUV transitions ($1s^2 2s$) in favour of (at least) $1s^2 2p$, which weakens the λλ1032, 1038 Å absorption but does not affect the $1s-2p$ X-ray absorption. However, it requires very high densities to maintain a finite excited state population, and, more directly, the measured wavelength of the X-ray line

is actually not consistent with the wavelength calculated for $1s-2p$ in excited O VI, off by about 0.03−0.05 Å, on the order of a full resolution element of both the *Chandra* LETGS and the *XMM-Newton* RGS (Raassen 2007, private communication). The conclusion is that the discrepancy is due to an authentic statistical fluctuation in the X-ray spectrum—or, more ironically, to the presence of a weak, slightly redshifted O VII absorption line.

5 Additional Aspects

5.1 Results from Numerical Simulations

Cosmological simulations not only have been used to investigate the large-scale distribution and physical state of the warm-hot intergalactic medium, they also have been applied to predict statistical properties of high-ion absorption systems that can be readily compared with the UV and X-ray measurements (e.g., Cen et al. 2001; Fang and Bryan 2001; Chen et al. 2003; Furlanetto et al. 2005; Tumlinson and Fang 2005; Cen and Fang 2006). Usually, a large number of artificial spectra along random sight-lines through the simulated volume are generated. Sometimes, instrumental properties of existing spectrographs and noise characteristics are modelled, too (e.g., Fangano et al. 2007). The most important quantities derived from artificial spectra that can be compared with observational data are the cumulative and differential number densities ($\mathrm{d}N/\mathrm{d}z$) of O VI, O VII, O VIII systems as a function of the absorption equivalent width. An example for this is shown in Fig. 8. Generally, there is a good match between the simulations and observations for the overall shape of the $\mathrm{d}N/\mathrm{d}z$ distribution (see also Sect. 3.2), but mild discrepancies exist at either low or high equivalent widths, depending on what simulation is used (see, e.g., Tripp et al. 2007). For the interpretation of such discrepancies it is important to keep in mind that the different simulations are based on different physical *models* for the gas, e.g., some simulations include galaxy feedback models, galactic wind models, non-equilibrium ionisation conditions, etc., others do not. For more information on numerical simulations of the WHIM see Bertone et al. (2008—Chap. 14, this issue).

WHIM simulations also have been used to investigate the frequency and nature of BLAs at low redshift (Richter et al. 2006b). As the simulations suggest, BLAs indeed host a substantial fraction of the baryons at $z = 0$. From the artificial UV spectra generated from their simulation Richter et al. derive a number of BLAs per unit redshift of $(\mathrm{d}N/\mathrm{d}z)_{\mathrm{BLA}} \approx 38$ for

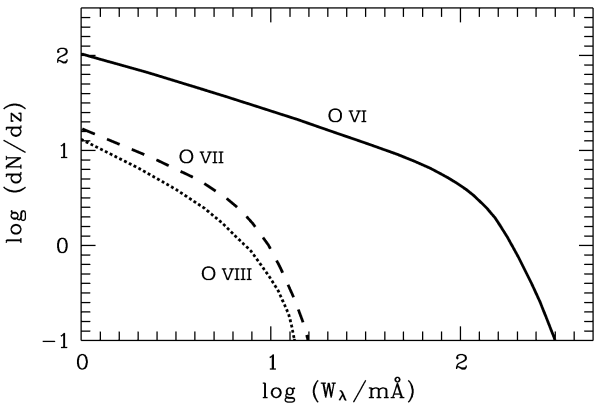


Fig. 8 The differential number of intervening oxygen high-ion (O VI, O VII,O VIII) absorbers in the WHIM in a cosmological simulation is plotted against the equivalent width of the absorption (for details see Cen and Fang 2006). While for O VII and O VIII no significant observational results are available to be compared with the simulated spectra (see Sect. 4.2), the predicted frequency of O VI absorbers is in good agreement with the observations (Sect. 3.2). Adapted from Cen and Fang (2006)

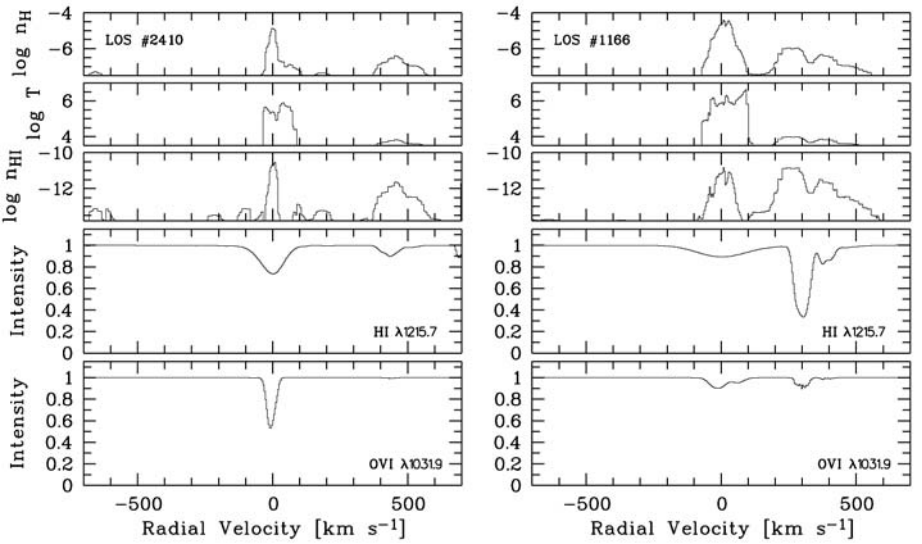


Fig. 9 Two examples for BLA absorbers from the WHIM in a cosmological simulation are shown. The panels show the logarithmic total hydrogen volume density, gas temperature, neutral hydrogen volume density, and normalised intensity for H I Lyα and O VI λ1031.9 absorption as a function of the radial velocity along each sightline. From Richter et al. (2006b)

H I absorbers with $\log(N(\mathrm{cm}^{-2})/b(\mathrm{km\,s}^{-1})) \geq 10.7$, $b \geq 40$ km s^{-1}, and total hydrogen column densities $N(\mathrm{H\,II}) \leq 10^{20.5}$ cm^{-2}. The baryon content of these systems is ~ 25 percent of the total baryon budget in the simulation. These results are roughly in line with the observations if partial photoionisation of BLAs is taken into account (Richter et al. 2006a; Lehner et al. 2007). From the simulation further follows that BLAs predominantly trace shock-heated collisionally ionised WHIM gas at temperatures $\log T \approx 4.4-6.2$. Yet, about 30 percent of the BLAs in the simulation originate in the photoionised Lyα forest ($\log T < 4.3$) and their large line widths are determined by non-thermal broadening effects such as unresolved velocity structure and macroscopic turbulence. Figure 9 shows two examples of the velocity profiles of BLAs generated from simulations presented in Richter et al. (2006b).

The results from the analysis of artificially generated UV spectra underline that the comparison between WHIM simulations and quasar absorption line studies indeed are quite important for improving both the physical models in cosmological simulations and the strategies for future observations of the warm-hot intergalactic gas.

5.2 WHIM Absorbers at High Redshift

Although this chapter concentrates on the properties of WHIM absorbers at low redshift (as visible in UV and X-ray absorption) a few words about high-ion absorption at high redshifts ($z > 2$) shall be given at this point. At redshifts $z > 2$, by far most of the baryons are residing in the photoionised intergalactic medium that gives rise to the Lyα forest. At this early epoch of the Universe, baryons situated in galaxies and in warm-hot intergalactic gas created by large-scale structure formation contribute together with only < 15 percent to the total baryon content of the Universe. Despite the relative unimportant role of the WHIM at high z, O VI absorbers are commonly found in optical spectra of high-redshift quasars (e.g., Bergeron et al. 2002; Carswell et al. 2002; Simcoe et al. 2004). The observation of intervening O VI

absorbers at high redshift is much easier than in the local Universe, since the absorption features are redshifted into the optical regime and thus are easily accessible with ground-based observatories. However, blending problems with the numerous H I Lyα forest lines at high z are much more severe than for low-redshift sightlines. Because of the higher intensity of the metagalactic UV background at high redshift it is expected that many of the O VI systems in the early Universe are photoionised. Collisional ionisation of O VI yet may be important for high-redshift absorbers that originate in galactic winds (see, e.g., Fangano et al. 2007). While for low redshifts the population of O VI absorbers is important for the search of the "mission baryons" that are locked in the WHIM phase in the local Universe, O VI absorbers at high redshift are believed to represent a solution for the problem of the "missing metals" in the early epochs of structure formation. This problem arises from the facts that at high redshift an IGM metallicity of ~ 0.04 is expected from the star-formation activity of Lyman-Break Galaxies (LBGs), while observations of intervening C IV absorption systems suggest an IGM abundance of only ~ 0.001 solar (Songaila 2001; Scannapieco et al. 2006), thus more than one order of magnitude too low. Possibly, most of the missing metals at high z are hidden in highly-ionised hot gaseous halos that surround the star-forming galaxies (e.g., Ferrara et al. 2005) and thus should be detectable only with high ions such as O VI rather than with intermediate ions such as C IV. Using the UVES spectrograph installed on the *Very Large Telescope* (VLT) Bergeron and Herbert-Fort (2005) have studied the properties of high-redshift O VI absorbers along ten QSO sightlines and have found possible evidence for such a scenario. Additional studies are required to investigate the nature of high-z O VI systems and their relation to galactic structures in more detail. However, from the existing measurements clearly follows that the study of high-ion absorbers at large redshifts is of great importance to our understanding of the formation and evolution of galactic structures at high z and the transport of metals into the IGM.

6 Concluding Remarks

The analysis of absorption features from high ions of heavy elements and neutral hydrogen currently represents the best method to study baryon content, physical properties, and distribution of the warm-hot intergalactic gas in large-scale filaments at low and high redshift. However, the interpretation of these spectral signatures in terms of WHIM baryon content and origin still is afflicted with rather large systematic uncertainties due to the limited data quality and the often poorly known physical conditions in WHIM absorbers (e.g., ionisation conditions, metal content, etc.). Future instruments in the UV (e.g., COS) and in the X-ray band (e.g., *XEUS*, *Constellation X*) hold the prospect of providing large amounts of new data on the WHIM with good signal-to-noise ratios and substantially improved absorber statistics. These missions therefore will be of great importance to improve our understanding of this important intergalactic gas phase.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialised X-ray components in clusters of galaxies". SRON is supported financially by NWO, the Netherlands Organisation for Scientific Research.

References

M. Asplund, N. Grevesse, A.J. Sauval, C. Allende Prieto, D. Kiselman, Astron. Astrophys. **417**, 751 (2004)
J. Bergeron, S. Herbert-Fort, in *Probing Galaxies through Quasar Absorption Lines*, ed. by P.R. Williams, C.-G. Shu, B. Menard. IAU Coll. 199 (Cambr. Univ. Press, 2005), p. 265

J. Bergeron, B. Aracil, P. Petitjean, C. Pichon, Astron. Astrophys. **396**, L11 (2002)
S. Bertone, J. Schaye, K. Dolag, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9318-3
D.V. Bowen, M. Pettini, J.C. Blades, Astrophys. J. **580**, 169 (2002)
J.N. Bregman, Annu. Rev. Astron. Astrophys. **45**, 221 (2007)
A.C. Brinkman, C.J.T. Gunsing, J.S. Kaastra et al., Astrophys. J. **530**, L111 (2000)
A.M. Bykov, F.B.S. Paerels, V. Petrosian, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9309-4
C.R. Canizares, J. Kruper, Astrophys. J. **278**, L99 (1984)
C.R. Canizares, G.W. Clark, T.H. Markert et al., Astrophys. J. **234**, L33 (1979)
C.R. Canizares, J.E. Davis, D. Dewey et al., Publ. Astron. Soc. Pac. **117**, 1144 (2005)
B. Carswell, J. Schaye, T.-S. Kim, Astrophys. J. **578**, 43 (2002)
R. Cen, T. Fang, Astrophys. J. **650**, 573 (2006)
R. Cen, J. Ostriker, Astrophys. J. **514**, 1 (1999)
R. Cen, T.M. Tripp, J.P. Ostriker, E.B. Jenkins, Astrophys. J. **559**, L5 (2001)
H.-W. Chen, J.X. Prochaska, Astrophys. J. **543**, L9 (2000)
X. Chen, D. Weinberg, N. Katz, R. Davé, Astrophys. J. **594**, 42 (2003)
C.W. Danforth, J.M. Shull, Astrophys. J. **624**, 555 (2005)
R. Davé, R. Cen, J. Ostriker et al., Astrophys. J. **552**, 473 (2001)
J.W.A. den Herder, A.C. Brinkman, S.M. Kahn et al., Astron. Astrophys. **365**, L7 (2001)
G.W. Drake, Can. J. Phys. **66**, 586 (1988)
T. Fang, G.L. Bryan, Astrophys. J. **561**, L31 (2001)
T. Fang, H.L. Marshall, G.L. Bryan, C.R. Canizares, Astrophys. J. **555**, 356 (2001)
T. Fang, H.L. Marshall, J.C. Lee, D.S. Davis, C.R. Canizares, Astrophys. J. **572**, L127 (2002)
T. Fang, K.R. Sembach, C.R. Canizares, Astrophys. J. **586**, L49 (2003)
T. Fang, R.A.C. Croft, W.T. Sanders et al., Astrophys. J. **623**, 612 (2005)
T. Fang, C.F. McKee, C.R. Canizares, M. Wolfire, Astrophys. J. **644**, 174 (2006)
A.P.M. Fangano, A. Ferrara, P. Richter, Mon. Not. R. Astron. Soc. **381**, 469 (2007)
A. Ferrara, E. Scannapieco, J. Bergeron, Astrophys. J. **634**, L37 (2005)
A.J. Fox, B.D. Savage, B.P. Wakker et al., Astrophys. J. **602**, 738 (2004)
R. Fujimoto, Y. Takei, T. Tamura et al., Publ. Astron. Soc. Jpn. **56**, L29 (2004)
S.R. Furlanetto, L.A. Phillips, M. Kamionkowski, Mon. Not. R. Astron. Soc. **359**, 295 (2005)
F. Haardt, P. Madau, Astrophys. J. **461**, 20 (1996)
W.R. Johnson, G. Soff, Atom. Data Nucl. Data Tables **33**, 405 (1985)
J.S. Kaastra, N. Werner, J.W.A. den Herder et al., Astrophys. J. **652**, 189 (2006)
J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-y
H. Kawahara, K. Yoshikawa, S. Sasaki et al., Publ. Astron. Soc. Jpn. **58**, 657 (2006)
R.A. Kimble, B.E. Woodgate, C.W. Bowers et al., Astrophys. J. **492**, L83 (1998)
N. Lehner, B.D. Savage, P. Richter et al., Astrophys. J. **658**, 680 (2007)
R. Lynds, Astrophys. J. **164**, L73 (1971)
S. Mathur, D.H. Weinberg, X. Chen, Astrophys. J. **582**, 82 (2003)
D. McCammon, R. Almy, E. Apodaca et al., Astrophys. J. **576**, 188 (2002)
H.W. Moos, W.C. Cash, L.L. Cowie et al., Astrophys. J. **538**, L1 (2000)
F. Nicastro, in *Maps of the Cosmos*, ed. by M. Colless, L. Staveley-Smith, R. Stathakis. IAU Symp. 216 (Astr. Soc. Pac., 2005), p. 297. astro-ph/0311162
F. Nicastro, A. Zezas, J. Drake et al., Astrophys. J. **573**, 157 (2002)
F. Nicastro, A. Zezas, M. Elvis et al., Nature **421**, 719 (2003)
F. Nicastro, S. Mathur, M. Elvis et al., Astrophys. J. **629**, 700 (2005)
W.R. Oegerle, T.M. Tripp, K.R. Sembach et al., Astrophys. J. **538**, L23 (2000)
F.B.S. Paerels, J.S. Kaastra, T. Ohashi et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9323-6
M. Pettini, S.L. Ellison, J. Bergeron, P. Petitjean, Astron. Astrophys. **391**, 21 (2002)
N. Prause, D. Reimers, C. Fechner, E. Janknecht, Astron. Astrophys. **470**, 67 (2007)
A.P. Rasmussen, S.M. Kahn, F.B.S. Paerels, in *The IGM/Galaxy Connection: The Distribution of Baryons at* $z = 0$, ed. by J.L. Rosenberg, M.E. Putman. ASSL Conf. Proc., vol. 281 (Kluwer, Dordrecht, 2003), p. 109. astro-ph/0301183
A.P. Rasmussen, S.M. Kahn, F. Paerels et al., Astrophys. J. **656**, 129 (2007)
P. Richter, Rev. Mod. Astron. **19**, 31 (2006c)
P. Richter, B.D. Savage, T.M. Tripp, K.R. Sembach, Astrophys. J. Suppl. Ser. **153**, 165 (2004)
P. Richter, B.D. Savage, K.R. Sembach, T.M. Tripp, Astron. Astrophys. **445**, 827 (2006a)
P. Richter, T. Fang, G.L. Bryan, Astron. Astrophys. **451**, 767 (2006b)
D.J. Sahnow, H.W. Moos, T.B. Ake et al., Astrophys. J. **538**, L7 (2000)
B.D. Savage, K.R. Sembach, E.B. Jenkins et al., Astrophys. J. **538**, L27 (2000)

B.D. Savage, K.R. Sembach, T.M. Tripp, P. Richter, Astrophys. J. **564**, 631 (2002)
B.D. Savage, N. Lehner, B.P. Wakker, T.M. Sembach, K.R. Tripp, Astrophys. J. **626**, 776 (2005)
B.D. Savage, K.R. Sembach, B.P. Wakker et al., Astrophys. J. Suppl. Ser. **146**, 125 (2003)
E. Scannapieco, C. Pichon, B. Aracil et al., Mon. Not. R. Astron. Soc. **365**, 615 (2006)
M. Schmidt, P. Beiersdorfer, H. Chen et al., Astrophys. J. **604**, 562 (2004)
K.R. Sembach, B.P. Wakker, B.D. Savage et al., Astrophys. J. Suppl. Ser. **146**, 165 (2003)
K.R. Sembach, T.M. Tripp, B.D. Savage, P. Richter, Astrophys. J. Suppl. Ser. **155**, 351 (2004)
J.M. Shull, S.V. Penton, J.T. Stocke et al., Astron. J. **116**, 2094 (1998)
R.A. Simcoe, W.L.W. Sargent, M. Rauch, Astrophys. J. **606**, 92 (2004)
A. Songaila, Astrophys. J. **561**, L153 (2001)
K. Steenbrugge, F. Nicastro, M. Elvis, in *The X-ray Universe 2005*, ed. by A. Wilson. ESA SP-604, 2006, p. 751. astro-ph/0511234
J.T. Stocke, J.M. Shull, S. Penton, M. Donahue, C. Carilli, Astrophys. J. **451**, 24 (1995)
J.T. Stocke, S.V. Penton, C.W. Danforth et al., Astrophys. J. **641**, 217 (2006)
R.S. Sutherland, M.A. Dopita, Astrophys. J. Suppl. Ser. **88**, 253 (1993)
Y. Takei, J.P. Henry, A. Finoguenov et al., Astrophys. J. **655**, 831 (2007)
T.M. Tripp, L. Lu, B.D. Savage, Astrophys. J. **508**, 200 (1998)
T.M. Tripp, B.D. Savage, E.B. Jenkins, Astrophys. J. **534**, L1 (2000)
T.M. Tripp, M.L. Giroux, J.T. Stocke, J. Tumlinson, W.R. Oegerle, Astrophys. J. **563**, 724 (2001)
T.M. Tripp, K.R. Sembach, D.V. Bowen et al., Astrophys. J. Suppl. Ser. (2007, in press). astro-ph/07061214
J. Tumlinson, T. Fang, Astrophys. J. **623**, L100 (2005)
P. Valageas, R. Schaeffer, J. Silk, Astron. Astrophys. **388**, 741 (2002)
D.A. Verner, E.M. Verner, G.J. Ferland, Atom. Data Nucl. Data Tables **64**, 1 (1996)
B.P. Wakker, B.D. Savage, K.R. Sembach et al., Astrophys. J. Suppl. Ser. **146**, 1 (2003)
Q.D. Wang, Y. Yao, T.M. Tripp et al., Astrophys. J. **635**, 386 (2005)
R.J. Williams, S. Mathur, F. Nicastro et al., Astrophys. J. **631**, 856 (2005)
R.J. Williams, S. Mathur, F. Nicastro, M. Elvis, Astrophys. J. **665**, 247 (2007)
B.E. Woodgate, Publ. Astron. Soc. Pac. **110**, 1183 (1998)

Chapter 4
Soft X-Ray and Extreme Ultraviolet Excess Emission from Clusters of Galaxies

F. Durret · J.S. Kaastra · J. Nevalainen · T. Ohashi · N. Werner

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9313-8

Abstract An excess over the extrapolation to the extreme ultraviolet and soft X-ray ranges of the thermal emission from the hot intracluster medium has been detected in a number of clusters of galaxies. We briefly present each of the satellites (EUVE, ROSAT PSPC and BeppoSAX, and presently XMM-Newton, Chandra and Suzaku) and their corresponding instrumental issues, which are responsible for the fact that this soft excess remains controversial in a number of cases. We then review the evidence for this soft X-ray excess and discuss the possible mechanisms (thermal and non-thermal) which could be responsible for this emission.

Keywords Galaxies: clusters · X-ray: spectra

F. Durret (✉)
Institut d'Astrophysique de Paris, CNRS, UMR 7095, Université Pierre et Marie Curie,
98bis Bd Arago, 75014 Paris, France
e-mail: durret@iap.fr

J.S. Kaastra · N. Werner
SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands

J.S. Kaastra
Astronomical Institute, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, The Netherlands

J. Nevalainen
Observatory, University of Helsinki, P.O. Box 14, 00014 Helsinki, Finland

T. Ohashi
Department of Physics, Tokyo Metropolitan University, 1-1 Minami-Osawa, Tokyo 192-0397, Japan

N. Werner
Max-Planck-Institute für Astrophysik, Karl-Schwarzschild-Strasse 1, 85749, Garching, Germany

1 Introduction

The existence of soft excess emission originating from clusters of galaxies, defined as emission detected below 1 keV in excess over the usual thermal emission from hot intracluster gas (hereafter the ICM) has been claimed since 1996. Soft excesses are particularly important to detect because they may (at least partly) be due to thermal emission from the Warm-Hot Intergalactic Medium, where as much as half of the baryons of the Universe could be. They are therefore of fundamental cosmological importance.

Soft excess emission has been observed (and has also given rise to controversy) in a number of clusters, mainly raising the following questions: (1) Do clusters really show a soft excess? (2) If so, from what spatial region(s) of the cluster does the soft excess originate? (3) Is this excess emission thermal, originating from warm-hot intergalactic gas (at temperatures of $\sim 10^6$ K), or non-thermal, in which case several emission mechanisms have been proposed. Interestingly, some of the non-thermal mechanisms suggested to account for soft excess emission can also explain the hard X-ray emission detected in some clusters, for example by RXTE and BeppoSAX (also see Petrosian et al. 2008—Chap. 10, this issue; Rephaeli et al. 2008—Chap. 5, this issue).

Several instruments have been used to search for soft excess emission: EUVE, ROSAT and BeppoSAX in the 1990s, and presently XMM-Newton, Chandra and Suzaku. We will briefly present a history of these detections, emphasising the difficulties to extract such weak signal and the underlying hypotheses, and summarising what is known on each of the observed clusters. For clarity, results will be presented separately for the various satellites. Finally we summarise and discuss the various models proposed to account for Extreme Ultraviolet (hereafter EUV) emission with their pros and cons.

2 Instrumental Issues

2.1 The *Extreme Ultraviolet Explorer* (EUVE) Satellite

The EUVE satellite was launched in 1992 and operated till 2001, covering the 50–250 eV energy range (Bowyer and Malina 1991). The rectangular shape of the Lex/B (65–248 eV) filter resulted in images where the length of one side of the image greatly exceeds the other side: $\sim 40'$ in width and more than 2° in length (see e.g. Fig. 1 in Durret et al. 2002).

The spatial scale was 13 pixels/arcmin. Due to the limited sensitivity of EUVE, exposure times on clusters were typically between several tens of ks and 1 Ms.

2.2 ROSAT PSPC

The Position Sensitive Proportional Counter PSPC instrument on board the ROSAT satellite first detected the soft excess component of clusters of galaxies at X-ray wavelengths. The PSPC had an effective area of about 200 cm^2 at 0.28 keV, enabling soft X-ray studies. It had a large field of view (hereafter FOV) of $50'$ radius, covering the virial radius in most clusters and enabling in most cases the estimation of the local background. The low and stable internal background of the PSPC allowed reliable X-ray measurements at large radii where the background is important. A major problem for soft excess studies with the PSPC is that it did not cover energies above 2 keV. Thus, it could not be used to determine reliably the hot gas properties, which had to be measured elsewhere. A further limitation was the low energy resolution which did not allow detection of possible emission line blends emanating from the soft component. The angular resolution of the ROSAT PSPC was $\sim 15''$. Details on the ROSAT PSPC instrument can be found in Briel et al. (1996).

2.3 XMM-Newton EPIC

An important change in the study of the soft excess came with the XMM-Newton satellite. The XMM-Newton European Photon Imaging Camera (EPIC) instruments PN and MOS extend the energy band coverage to 10 keV, thus enabling simultaneous determination of the hot gas and soft excess component properties.

The large collecting area of the EPIC telescopes ($\sim$1000 cm^2 at 0.5 keV for the PN) provides the high statistical quality data necessary to examine the few 10% soft excess effect on top of the hot gas emission. The spectral resolution of the PN at 0.5 keV is 60 eV (FWHM), rendering it possible to resolve the emission lines emanating from the soft excess component. However, the relatively small FOV (15′ radius) of EPIC prevents the study of the cluster outskirts for the nearest clusters with single pointings. The usage of offset pointings introduce the complex stray light problem which complicates the analysis of weak signals such as the soft excess. The more distant clusters, which would be covered out to the virial radius with a single pointing, are fainter, which reduces the quality of the signal. Thus, the XMM-Newton soft excess analysis is mostly limited to the central regions of nearby clusters. Details on the characteristics of XMM-Newton can be found in Turner et al. (2001) and Strüder and Briel (2001), or in Ehle et al. (2006).

A further complication is the strong and flaring particle-induced detector background. A local background estimate is vital when analysing weak signals such as the cluster soft excess. This is usually not available for nearby clusters since they fill the FOV, and one has to resort to blank-sky based background estimates. This introduces uncertainties in the analysis, and further limits the cluster analysis to central regions where the background is not important. Together with the FOV limitations, the background problem limits the usefulness of XMM-Newton for measuring the soft excess in a large cluster sample.

The on-going calibration work on the PN and MOS instruments resulted in changes in the derived soft excess properties for a few clusters (Nevalainen et al. 2007). Thus, the EPIC results have some degree of systematic uncertainty based on calibration inaccuracies, and definitive results on the soft excess properties are not yet available (also see Sect. 4.1.2).

2.4 Chandra

Chandra, launched in the same year as XMM-Newton (1999), has a similar eccentric orbit as the latter satellite and therefore suffers from comparable enhanced background problems. Its angular resolution is much higher (0.8″ Half Energy Width) than that of XMM-Newton (14″). The advantage is that for extended sources like clusters of galaxies subtraction of contaminating background point sources can be done more accurately, and more importantly, blurring effects by the point spread function of the telescope can usually be ignored. These blurring effects were a serious problem for the BeppoSAX LECS data (Sect. 3.2.2). However, the effective area at low energies ($E < 0.5$ keV) of Chandra is an order of magnitude smaller than that of XMM-Newton, and time-dependent contamination that affects in particular the lower energies is a complicating factor in the analysis of Chandra data.

2.5 Suzaku XIS

Suzaku is the fifth Japanese X-ray astronomy satellite, launched in July 2005 (Mitsuda et al. 2007). Unfortunately, the main observing instrument XRS, consisting of X-ray microcalorimeters used in space for the first time, did not last until the first space observations,

due to the loss of liquid He (Kelley et al. 2007). On the other hand, the XIS instrument, which employs X-ray CCDs with improved performance for X-ray spectroscopy, is functioning well (Koyama et al. 2007). Regarding soft X-ray spectral studies of diffuse sources, XIS offers the best capability so far achieved with X-ray satellites.

The XIS system consists of 4 units of mirror and detector combinations. The X-ray mirrors are identical and have a focal length of 4.5 m and an effective area of about 500 cm^2 at 2 keV (Serlemitsos et al. 2007). The angular resolution is about $2'$, limited by the lightweight design of the thin foil mirror.

The 4 CCDs in the focal plane operate jointly during the observations. The chips are square with an area of 10 mm $\times$ 10 mm and the number of pixels is 10^6. The field of view is $17' \times 17'$. One chip is back-illuminated, which gives a superior soft X-ray sensitivity with somewhat poorer energy resolution and higher background above 7 keV. The other 3 chips are of standard front illumination type. The typical energy resolution was 150 eV FWHM at 5.9 keV at the time of launch. The resolution degraded significantly with time (200 eV after 1 year), and the XIS team performed a charge injection operation after October 2006 to maintain the resolution around 170 eV. In November 2006, one of the 3 FI (front illumination) chips developed excess noise, and it has been switched off since then. This leaves a total of 3 chips, one BI (back illumination) and 2 FIs, in operation. Its background is lower than that of XMM-Newton.

3 Soft X-Ray Excess Emission Based on EUVE and ROSAT PSPC Data

3.1 The First Objects with a Soft X-Ray Excess Discovered: Virgo (Redshift $z = 0.0038$) and Coma ($z = 0.0231$)

The first mention of soft excess emission in the 0.065–0.245 keV band was made by Lieu et al. (1996a) for the Virgo cluster, based on data obtained with the EUVE satellite. These authors detected emission around M 87 up to a radial distance of $\sim 20'$, and found excess emission over the best fit single temperature plasma model obtained by fitting simultaneously the EUVE and ROSAT PSPC 0.18–2.0 keV data. This excess was found to decrease with radial distance (see Fig. 1). Similar results were then found for the Coma cluster, also based on EUVE and ROSAT data, by Lieu et al. (1996b).

The first interpretation proposed was that the soft excess emission was thermal, with a single component of ICM gas at a temperature between 5×10^5 K and 10^6 K for Virgo (Lieu et al. 1996a) and two components at 8×10^5 K and 2×10^6 K for Coma (Lieu et al. 1996b). The immediate implication of this hypothesis was that such warm gas would cool very rapidly, implying a very high mass accretion rate of several hundred Solar masses per year, and therefore the production of a large amount of gas in the cool phase (of the order of 10^{14} $M_\odot$, Lieu et al. 1996a, 1996b).

A second model was then proposed by Sarazin and Lieu (1998), in which soft excess emission was due to the Inverse Compton emission (hereafter IC) of cosmic microwave background photons on a relic population of cosmic ray electrons. Such electrons could have been produced by supernovae, by radio galaxies or by particle acceleration in intracluster shocks. Their energy would amount to 1–10% of the thermal energy content of the ICM. The inverse Compton model was also favoured by Bowyer and Berghöfer (1998), and Lieu et al. (1999a) suggested that cosmic rays could be as energetically important as the thermal ICM.

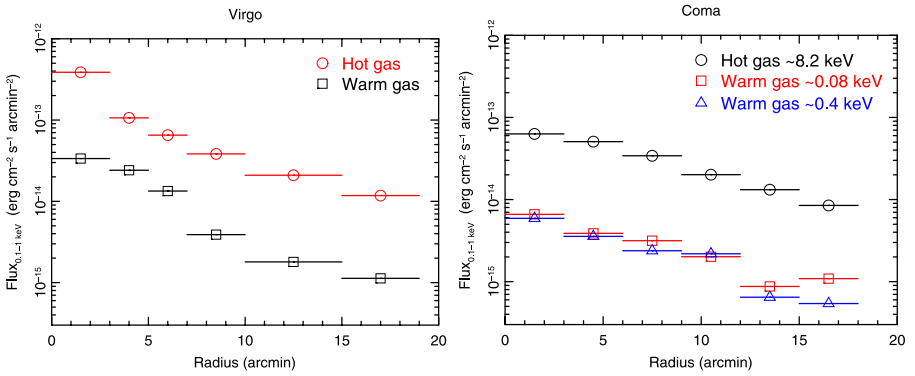


Fig. 1 *Left panel*: Radial profiles of the 0.1–1.0 keV surface brightness of the hot ICM and of the warm gas in Virgo found by Lieu et al. (1996a). *Right panel*: Radial profile of the 0.1–1.0 keV surface brightness of the hot ICM and of the two warm gas components in Coma found by Lieu et al. (1996b)

From ROSAT images at low energy, Bonamente et al. (2001a) derived the existence of large amounts of cold gas in the ICM of Virgo and Coma, confirming the existence of a multiphase ICM.

More clusters were then observed, Virgo and Coma were reobserved, and controversy began to arise concerning the very existence of a soft excess over the thermal emission from hot gas responsible for cluster X-ray emission.

An agreement on the existence of soft excess emission based on EUVE and ROSAT PSPC data was reached by different observers only on the Virgo and Coma clusters. By taking into account the variations of telescope sensitivity over the field of view, Bowyer et al. (1999) confirmed the presence of EUV excess emission in Coma. These authors also later confirmed the soft excess in Virgo and interpreted it as due to inverse Compton radiation (Berghöfer et al. 2000a). With EUVE *in situ* background measurements, Bonamente et al. (2001b) agreed with the existence of an EUV excess in Virgo and found that this excess increased with radius.

Using PSPC data, Bonamente et al. (2003) found a very large scale (2.6 Mpc) soft excess halo in the Coma cluster, exceeding the thermal emission by 200% at the largest radii. Note that at a distance of 40′ the soft excess flux exceeds that of the hot gas by 100%, consistent with the later XMM-Newton analysis of Finoguenov et al. (2003) at the same distance. The non-thermal model gave a poor fit to the spectrum of the soft excess, while the thermal model was acceptable, with a best-fit temperature of ∼0.2 keV.

3.2 Other Clusters Observed with EUVE and ROSAT

3.2.1 Abell 1795 ($z = 0.0625$)

A strong soft excess was found in Abell 1795 from EUVE data by Mittaz et al. (1998). Later, Bonamente et al. (2001b) confirmed the existence of this excess by observing with EUVE an *in situ* background, and found that, as for Virgo, the EUV excess increased with radius.

However, these results were debated: by taking into account the variations of telescope sensitivity over the field of view, Bowyer et al. (1999) claimed that there was no evidence for EUV excess in Abell 1795 and later confirmed this result from BeppoSAX data (Berghöfer and Bowyer 2002). However, the existence of soft excess emission in this cluster was confirmed by Durret et al. (2002) from EUVE and ROSAT PSPC data. Soft excess in Abell 1795

was also found with the ROSAT PSPC (Bonamente et al. 2003) at the 10–20% level in the 0.2–0.4 keV band, and by Nevalainen et al. (2003) and Kaastra et al. (2003), both based on XMM-Newton data (see Sects. 4.2.1 and 4.2.2, respectively).

3.2.2 Abell 2199 ($z = 0.0301$)

The presence of a soft excess was disputed in Abell 2199. From EUVE data, Lieu et al. (1999b) detected a soft excess, and confirmed its existence up to a radius of $20'$ using a background measured from offset pointing Lieu et al. (1999c). Confirmation of the existence of a soft and a hard excess in this cluster was given by Kaastra et al. (1999) based on BeppoSAX, EUVE and ROSAT data. A three phase ICM was proposed to account for the observations by Lieu et al. (2000).

On the other hand, from other EUVE observations taking into account the variations of telescope sensitivity over the field of view, Bowyer et al. (1999) found no evidence for EUV excess. They confirmed this result with BeppoSAX data (Berghöfer and Bowyer 2002), but Kaastra et al. (2002) claimed that their analysis was wrong. According to the latter authors, the problem was that Berghöfer & Bowyer based their conclusion on a plot of the ratio of the observed radial intensity profiles in the 0.5–2.2 keV band as compared to the 0.1–0.3 keV band. Because the point spread function of the BeppoSAX LECS instrument is a strong function of energy ($10'$ FWHM at $E = 0.28$ keV, scaling as $E^{-0.5}$), and the low-energy counts in the 0.1–0.3 keV band contain a significant contribution from higher energy photons, even for a fully isothermal cluster the ratio of the radial intensities in both bands shows strong ($\sim$50%) excursions. Moreover, the predicted ratio depends strongly on issues like abundance gradients, modelling of the cooling flow, etc. that were ignored by Berghöfer and Bowyer (2002).

The existence of a soft excess was independently confirmed by Durret et al. (2002) from EUVE and ROSAT PSPC data. The ROSAT PSPC analysis of Bonamente et al. (2003) also revealed marginal evidence for soft excess in Abell 2199, and they derived that the temperature of the soft excess component must be below 0.2 keV. Kaastra et al. (2003) confirmed the presence of a soft excess in this cluster from XMM-Newton data.

3.2.3 Abell 4059 ($z = 0.0460$)

No soft excess was detected in Abell 4059 by Berghöfer et al. (2000b); strangely, these authors even found a deficit of EUV emission in the central $2'$. However, a soft excess was detected in this cluster by Durret et al. (2002) as described in the next subsection.

3.2.4 Complementary Results on the Five above Clusters

A different approach was proposed by Durret et al. (2002) based on the wavelet analysis and reconstruction of EUVE and ROSAT PSPC observations for the five previously discussed clusters: Virgo, Coma, Abell 1795, Abell 2199 and Abell 4059. A soft excess was found in all five clusters, even when taking into account temperature and abundance gradients of the ICM. The radial profiles of the EUV to X-ray ratios are shown in Fig. 2. The EUV and X-ray profiles were shown to differ statistically, suggesting that the EUV and X-ray emissions were probably not due to the same physical mechanism.

3.2.5 Fornax ($z = 0.0046$)

No soft excess was detected in Fornax by Bowyer et al. (2001) at any scale, but Bonamente et al. (2002) did detect a soft X-ray excess in this cluster, which is part of the sample described in Sect. 3.2.9.

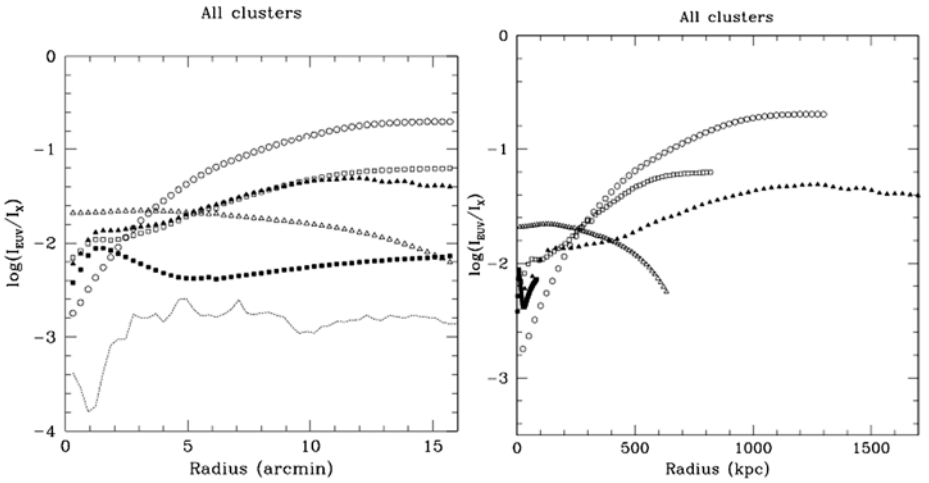


Fig. 2 EUV to X-ray ratio as a function of cluster radius (expressed in arcmin in the *left plot* and in kpc in the *right plot*) for the five clusters observed by Durret et al. (2002) with the EUVE and ROSAT PSPC satellites. The symbols are the following: Abell 1795: *filled triangles*, Abell 2199: *empty squares*, Abell 4059: *empty circles*, Coma: *empty triangles*, Virgo: *filled squares*. The *dashed line* in the *left figure* shows the ratio of the EUVE to ROSAT PSPC PSFs. Error bars are omitted for clarity; they are typically smaller than ± 0.01 in logarithmic scale, i.e. too small to be clearly visible on the figure

3.2.6 Shapley Supercluster ($z = 0.046$)

The ROSAT PSPC data of four clusters Abell 3571, Abell 3558, Abell 3560 and Abell 3562 in the Shapley supercluster were analysed by Bonamente et al. (2001c). They modelled the hot gas spectra using published ASCA results (Abell 3571 and Abell 3558) or fitting the PSPC data (Abell 3560 and Abell 3562). These models fit the PSPC data poorly due to excess residuals in the R2 band (0.15–0.3 keV) in all clusters. Based on the residuals, the authors determined the fractional soft excess emission above the hot gas emission model, amounting to 50% at most. For Abell 3571 and Abell 3558, they found that the fractional soft excess increased with the cluster radius, an effect already found in the EUVE data of Virgo, Abell 1795, and Abell 2199 (Bonamente et al. 2001b; Lieu et al. 1999b).

BeppoSAX LECS data on Abell 3571 and Abell 3562 supported the PSPC results. The analysis showed that a simple one-temperature thin-plasma model is not adequate to describe the emission.

The data indicated a thermal model with $kT = 0.1$ keV for Abell 3558, but the quality of the data was not appropriate for distinguishing the nature of the soft component or modelling its properties in detail.

3.2.7 Sérsic 159-03 ($z = 0.0580$)

Sérsic 159-03 (also known as AS 1101) was found to show strong soft excess and even to be the brightest soft excess cluster by Bonamente et al. (2001d). These authors combined three PSPC pointings of Sérsic 159-03 obtaining 19 ks of exposure time. The low temperature ($\sim$2 keV) of the hot gas allowed its spectroscopic analysis using the PSPC data alone. Extrapolating the best-fit spectral model from the 0.5–2.0 keV band to lower energies revealed a very strong soft excess at 0.2 keV, 100% over the hot gas emission. This excess was then confirmed with XMM-Newton (see Sect. 4.1.3).

Bonamente et al. (2001d) tested the hypothesis that the soft excess could be an artefact caused by a sub-Galactic absorbing column density $N_{\rm H}$ towards the cluster. However, this yielded an unrealistically low $N_{\rm H}$ (5×10^{19} cm^{-2}), even lower than that of the global minimum of the Lockman Hole, implying that $N_{\rm H}$ variation is not the cause of the soft excess in this cluster.

Despite the high quality of the signal, the nature of the soft excess component could not be determined, since non-thermal and thermal models yield statistically equally acceptable fits.

3.2.8 The Hercules Supercluster ($z \sim 0.035$)

A soft excess was also detected from low energy ROSAT data in the clusters Abell 2052 ($z = 0.0348$) and Abell 2063 ($z = 0.0348$) of the Hercules supercluster (Bonamente et al. 2005b); however these authors underlined the difficulty of background subtraction.

3.2.9 The Bonamente et al. (2002) Sample

Bonamente et al. (2002) published a ROSAT PSPC analysis of the largest soft excess sample of galaxy clusters so far, 38 in number. They modelled the ROSAT PSPC data of the hot gas using published results from ASCA and BeppoSAX which are more suitable due to their energy coverage up to 10 keV. These models were extrapolated to the PSPC channels and compared with the data in the 0.2–0.4 keV band. The analysis yielded significant (3σ) soft excess in ~30% of the clusters in the sample. The actual fraction of clusters with soft excess may be higher, since the clusters with the deepest ROSAT observations all had soft excess. Using a thermal model with a temperature of 0.08 keV and metal abundance of 0.3 Solar, or a power-law $I = I_0 E^{-\alpha_{\rm ph}}$ (where I is the intensity and E the photon energy) of index $\alpha_{\rm ph} = 1.75$, the soft excess fluxes correspond to 0.2–0.4 keV luminosities of 10^{40}–10^{43} erg s^{-1}. The excess was more prominent in the regions outside the central 150–200 kpc.

The indicated increase of the relative strength of the soft excess with radius is qualitatively consistent with the non-thermal Inverse Compton radiation of CMB photons from the relativistic cluster electrons. The hot gas emission is proportional to $n^2_{\rm e,thermal}$ while the IC is proportional to $n_{\rm e,relativistic}$ and the usual assumption is that the relativistic electrons follow the gas distribution. This model however faces energetic problems.

The authors also examined the possibility that the soft excess originates from a large number of unresolved X-ray-emitting cluster galaxies. The most relevant sources of X-rays are Low Mass X-ray Binaries (LMXB) and especially their k$T \sim 0.25$ keV blackbody component coming from the accretion disk. The typical soft X-ray luminosity of a galaxy due to LMXB is of order 10^{38} erg s^{-1} and thus 10^3–10^4 galaxies are needed to produce the typical observed soft X-ray luminosity, while typical numbers of galaxies in rich clusters are of order several hundreds. Thus, unresolved sources in cluster galaxies cannot explain the soft excess. A similar conclusion can be reached by looking at the galaxy X-ray luminosity function, such as that obtained for the Coma cluster by Finoguenov et al. (2004); in view of the relatively faint X-ray luminosities of the individual galaxies, the number of X-ray emitting galaxies required to account for the soft excess would be unrealistically high.

4 Soft X-Ray Emission Based on XMM-Newton EPIC Data

4.1 Continuum Detections

4.1.1 The Nevalainen et al. (2003) *Sample*

Nevalainen et al. (2003) found evidence for soft X-ray excess in clusters of galaxies using XMM-Newton EPIC data. They derived a 20–40% soft excess in the central 500 kpc regions of the clusters Coma, Abell 1795 and Abell 3112 ($z = 0.0703$) in channels below 2 keV, consistently in both PN and MOS instruments, and in ROSAT PSPC data.

A thermal model fits the data better than a non-thermal one, but at the level of calibration accuracy at the time, the non-thermal nature of the soft excess could not be ruled out.

Thermal modelling yielded temperatures in the range of 0.6–1.3 keV and metal abundances consistent with zero. Assuming that this warm gas occupies the same volume as the hot gas, the electron densities are of the order of 10^{-4}–10^{-3} cm^{-3}. These values lead to a cooling time scale larger than the Hubble time, i.e. the structures are self-consistent.

4.1.2 The Kaastra et al. (2003) *Sample*

Kaastra et al. (2003) examined a sample of 14 clusters of galaxies observed with XMM-Newton, and found significant evidence for soft X-ray excess continuum emission in 5 of them: Coma (centre studied only, since it is closer than the other 4), Abell 1795, Sérsic 159-03, Abell 2052 and MKW 3s. Different modelling of the hot gas temperature and metal abundance yielded lower temperatures for the soft component ($\sim$0.2 keV), compared to Nevalainen et al. (2003). The surface brightness of the warm gas is rather constant with radius, while that of the hot gas decreases with radius, falling below the warm gas surface brightness between 0.5 and 1 Mpc from the cluster centre. Note that in Coma, the central pointing does not cover this radius, so the hot gas brightness remains above the cool gas brightness, consistent with the above number. Such a behaviour is consistent with the Warm Hot Intergalactic Medium (hereafter WHIM) filament scenario, whereby the projected external filamentary structure is more extended than the cluster. Later, Kaastra et al. (2004) extended this sample to 21 clusters, and found 7 objects with soft excesses, the new cases being Abell 3112 (also in Nevalainen's paper) and Abell 2199.

Based on the quality of the spectral fits, in most cases the non-thermal model was also acceptable. Fitting the soft excess with a power-law model yielded rather constant photon indices ($\sim$2) with radius in a given cluster. At the cluster centre the luminosity of the non-thermal component is $\sim$10% of that of the hot gas, while the percentage increases towards 100% at the largest radii.

Bregman and Lloyd-Davies (2006) challenged the results of Kaastra et al. (2003), arguing that their soft excess detection was due to incorrect background subtraction. However, in a rebuttal paper, Nevalainen et al. (2007) showed that, especially in the central regions, the cluster emission is so bright compared to the background, that the details of the background modelling are insignificant. Thus, Bregman & Lloyd-Davies's claim appears to be unjustified.

Nevalainen et al. (2007) also found that the changes in the EPIC calibration between the years 2002 and 2005 resulted in a decrease of the soft excess signal. Using the Kaastra et al. (2003) modelling, the PN soft excess even disappeared in some clusters. However, the MOS instrument still detects a soft excess in all the clusters of Kaastra's sample. A less conservative hot gas modelling (Nevalainen et al., in preparation) with the current calibration information obtains consistent soft excesses in both PN and MOS data.

4.1.3 Sérsic 159-03

One of the first clusters to be observed by XMM-Newton was Sérsic 159-03. This observation with a net exposure time of 30 ks was taken in 2000 (Kaastra et al. 2001). The same observation was used for a search for soft excess emission (Kaastra et al. 2003, see Sect. 4.1.2 above). A 60 ks observation taken two years later was analysed by Bonamente et al. (2005a) and de Plaa et al. (2006).

Bonamente et al. (2005a) confirmed the existence of the soft X-ray excess emission in Sérsic 159-03 out to a distance of 1 Mpc from the cluster centre. The soft excess in the 0.3–1.0 keV band increases from 10% at the centre to 80% at the largest radii. The properties of the soft excess differ from those derived for Sérsic 159-03 using PSPC data (Bonamente et al. 2001d), likely due to different modelling of the data: XMM-Newton allows to determine the hot gas component unambiguously in the 2.0–7.0 keV band where the soft component has a negligible contribution.

The same dataset of Sérsic 159-03 was also analysed by de Plaa et al. (2006). However, they focused their attention on the chemical evolution of the cluster and they determined the background for the spectral analysis in the $9'$–$12'$ region around the cluster, subtracting the spatially extended soft excess emission together with the soft foreground emission. For the remaining soft excess in the cluster core they propose a non-thermal emission mechanism arising from IC scattering between CMB photons and relativistic electrons accelerated in bow shocks associated with ram pressure stripping of infalling galaxies.

Werner et al. (2007) studied Suzaku data together with the two XMM-Newton data sets of Sérsic 159-03 obtained two years apart mentioned before. They found consistent soft excess fluxes with all instruments in all observations. From the XMM-Newton data they derived radial profiles and 2D maps that show that the soft excess emission has a strong peak at the position of the central cD galaxy and does not show any significant azimuthal variations. They concluded that the spatial distribution of the soft excess is neither consistent with the models of intercluster warm-hot filaments, nor with models of clumpy warm intracluster gas associated with infalling groups as proposed by Bonamente et al. (2005a). Using the data obtained with the XMM-Newton RGS, Werner et al. could not confirm the presence of warm gas in the cluster centre with the expected properties assuming the soft excess was of thermal origin. They therefore concluded that the soft excess in Sérsic 159-03 is most probably of non-thermal origin.

4.2 Line Emission Detection

A crucial piece of evidence for the thermal nature of the soft X-ray excess would be the detection of emission lines. At the indicated temperatures of 0.1–0.5 keV the most prominent emission lines are from O VII and O VIII. The resonance, intercombination and forbidden lines of O VII have energies of 574, 569 and 561 eV (see also Kaastra et al. 2008—Chap. 9, this issue). For a low density plasma at 0.2 keV temperature, the centroid of the triplet has an energy of 568.7 eV (see Kaastra et al. 2003).

Geocoronal and heliospheric solar wind charge exchange (Wargelin et al. 2004; Fujimoto et al. 2007) also produces soft X-ray emission lines. These lines can vary by a factor of 3 on time scales of hours, which makes them difficult to model and subtract. The maximum observed brightness of these lines in the 0.5–0.9 keV range can reach the level of the cosmic X-ray background. Therefore, when the background level is important for the analysis, care should be taken when estimating the contaminating effects of this emission.

Finoguenov et al. (2003) observed a spatial variation of the soft excess emission in the Coma cluster. In the Coma 11 field they found an excess emission which is particularly

strong. Recent observations with the Suzaku satellite (Takei et al., private communication) do not confirm the level of O VII and O VIII line emission reported by Finoguenov et al. (2003) for the Coma 11 field, but they are consistent with the lowest reported values for the other fields in Coma observed with XMM-Newton. They suggest that a large fraction of the reported excess soft X-ray emission and the line emission observed in the Coma 11 field with XMM-Newton was due to Solar wind charge exchange emission.

4.2.1 The Kaastra et al. (2003) Sample

Kaastra et al. (2003) found evidence for O VII line emission in the clusters Sérsic 159-03, MKW 3s ($z = 0.0450$) and Abell 2052 in form of line-like residuals on top of the hot gas model at wavelengths consistent with the cluster redshifts. Note that the line emission is significant only in the outer regions (4′–12′) of the clusters.

The uncertainties also allow a Galactic origin ($z = 0$) for the emission. Also, Nevalainen et al. (2007) pointed out that a simple cluster-to-background flux level comparison does not exclude that the reported O VII line emission in the 0.5–0.65 keV band contains contributions from the geocoronal and heliospheric Solar wind charge exchange (see above). This possibility remains to be studied in detail.

Therefore this study does not give conclusive evidence for O VII line emission in these three clusters.

4.2.2 Coma

Finoguenov et al. (2003) analysed XMM-Newton data in the outskirts of the Coma cluster. The spectrum at ∼1 Mpc distance from the centre exhibits a very strong, ∼100% fractional soft excess continuum in channels below 0.8 keV, and the authors detect two emission lines in the 0.5–0.6 keV band. Both the excess continuum and the line features are well fit with a thermal model of $kT = 0.22$ keV (see Fig. 3). The temperature, as well as the derived baryon overdensity for the warm gas (∼200) are consistent with the WHIM filament properties in numerical simulations (e.g. Davé et al. 2001; Yoshikawa and Sasaki 2006). Let us note however that the soft excess in Finoguenov's model consists mostly of line emission, but due to the low spectral resolution of EPIC at low energy, it looks more like a continuum.

The redshift of the O VII line (0.0–0.026) is consistent with that of Coma (0.023) and also with being Galactic (0.0), implying that the emitter is located between us and Coma. Optical data show that there is a significant galaxy concentration in the direction of the warm gas in front of Coma. This suggests that the warm gas and the galaxy structure are connected, consistent with the WHIM filament scenario.

However, since the soft excess flux in Coma is comparable to the background level, the charge transfer flux in an active state (i.e. at the maximum observed level) can reach the soft excess level detected in Coma (Finoguenov et al. 2003). Bowyer and Vikhlinin (2004) use this coincidence as a proof that the soft excess detection in Coma is due to charge transfer. However, they did not model the soft excess spectrum with a charge transfer emission model to show that the model is consistent with the observed spectral features in Coma. Thus, their claim of Coma soft excess being due to charge transfer mechanism is not proven.

An indication for the presence of WHIM in Coma was also recently reported by Takei et al. (2007b) at 2.3σ level in absorption and in emission, based on RGS spectra of the Active Galactic Nucleus X–Comae, which is located behind the Coma cluster.

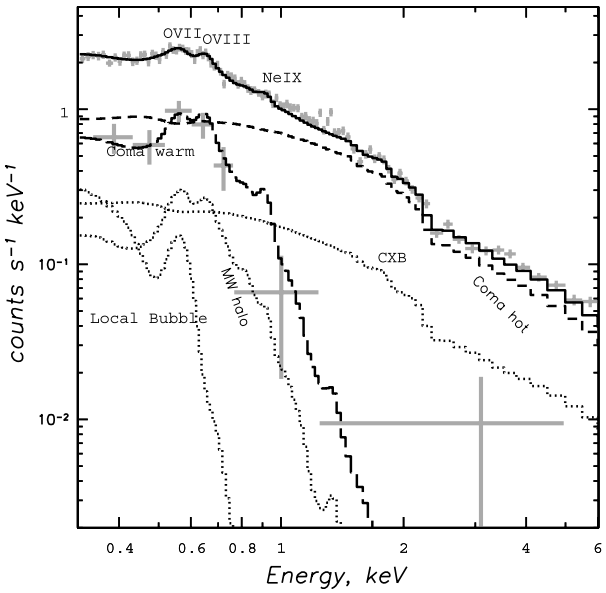


Fig. 3 Different emission components of the Coma spectrum obtained with the XMM-Newton PN instrument (from Finoguenov et al. 2003)

5 Soft X-Ray Emission Based on Suzaku Observations

5.1 Search for the Soft Excess in and around Clusters

Suzaku has observed about 30 clusters of galaxies by April 2007. Several of these observations were done in order to search for soft excess emission. We describe here the results for Abell 2218 and Abell 2052 in some detail since these two clusters have been studied rather extensively, and then briefly review results from other sources.

5.2 Abell 2218 ($z = 0.1756$)

Abell 2218 is a bright cluster at $z = 0.1756$ with an ICM temperature of about 7 keV. The Suzaku observations were carried out on two occasions with a total exposure of 80 ks and the results were published by Takei et al. (2007a). Several conditions are favourable to the detection of a soft excess in this cluster with Suzaku. First, the cluster is known to be undergoing a merger, likely along the line of sight, as indicated by two galaxy concentrations at different redshifts. This picture is supported by X-ray observations from Chandra and by the presence of remarkable arcs due to gravitational lensing. Such a merger suggests that a large-scale filament may be located along the line of sight and that the column density of oxygen contained in the warm filament gas could be high. The second reason is that the XIS energy resolution allows to separate a $\sim$100 eV cosmologically redshifted oxygen line from a zero redshift oxygen line. Background data in two regions at about 5° offset from the cluster were also taken to obtain information on the oxygen line intensity of the Galactic component.

The time variation of the soft X-ray flux during the observations was very small, so the effect from the solar wind charge exchange emission was considered to be small. A soft excess was searched for in the outside region of the cluster, i.e. for the region with radius greater than 5′ (880 kpc at the source) from the Abell 2218 centre. The observed spectra

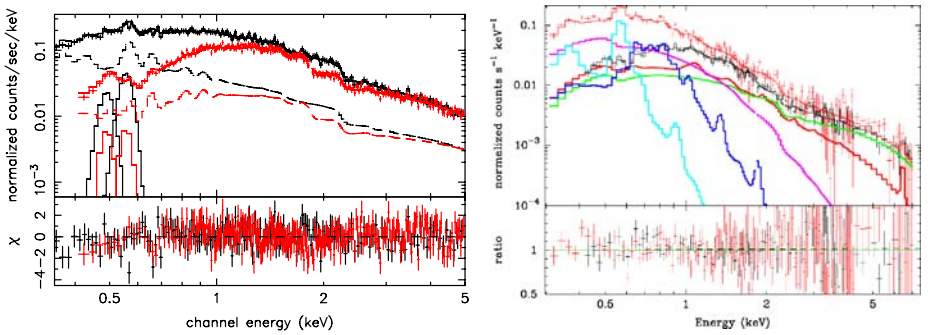


Fig. 4 *Left*: Observed XIS spectrum for the outer region of Abell 2218, with upper limits for the redshifted O VII and O VIII lines, for the BI (*black*) and FI (*red*) sensors. The background spectra are shown with *dashed lines*. This figure is taken from Takei et al. (2007a). *Right*: Observed XIS spectrum for $r = 12'$–$15'$. Spectral fit including Galactic emission (0.1 and 0.3 keV in *light* and *dark blue*), cosmic X-ray background (*green*), and thermal ICM emission with k$T = 2.8$ keV (*red*) requires an additional thermal Bremsstrahlung component with k$T = 0.7$ keV (*magenta*) whose intensity is constant over $r = 0$–$20'$. The summed up model (*orange*) is fitted to the data. Observations with Suzaku confirm the presence of soft excess emission in Sérsic 159-03 and its derived flux is consistent with the values determined with XMM-Newton (Sect. 4.1.3, Werner et al. 2007). However, Suzaku does not confirm the presence of the redshifted O VII lines in the cluster. The excess emission can be fit statistically equally well with a thermal model with low oxygen abundance (< 0.15 Solar) and with a non-thermal model. The figure is taken from Tamura et al. (2007)

(see Fig. 4) show structures around 0.5 and 0.6 keV, which are the energies corresponding to redshifted O VII and O VIII lines. However, close examination showed that the hump at 0.5 keV can be caused by the oxygen edge in the XIS filters, and the peak around 0.6 keV can be due to Galactic O VII line emission. This unfortunate situation hampered the original XIS capability of detecting the soft excess associated with the cluster. Takei et al. (2007a) set upper limits for the O VII and O VIII line intensities as shown in their Fig. 1, including all the uncertainties in the instrumental response (e.g. by the enhanced contamination in the XIS filters) and in the Galactic line intensities.

Even though the observing conditions were not optimal, the upper limits obtained are nearly an order of magnitude lower than the soft excess level reported in other clusters by Kaastra et al. (2003). With this upper limit, the density of the warm gas can be constrained. The limit for the gas density is:

$$\delta \equiv n_{\rm H}/\bar{n}_{\rm H} < 270\left(\frac{Z}{0.1\ Z_\odot}\right)^{-1/2}\left(\frac{L}{2\ {\rm Mpc}}\right)^{-1/2}, \qquad (1)$$

where $\bar{n}_{\rm H} = X\Omega_{\rm b}\rho_{\rm crit}(1+z)^3/m_{\rm p} = 1.77\times10^{-7}(1+z)^3\ {\rm cm}^{-3}$ is the mean hydrogen density in the universe, where $X = 0.71$ is the hydrogen to total baryon mass ratio, $\Omega_{\rm b} = 0.0457$ is the baryon density of the universe, $\rho_{\rm crit} = 9.21\times10^{-30}\ {\rm g\,cm}^{-3}$ is the critical density of the universe, and $m_{\rm p}$ is the proton mass. Even though this level of gas density is much higher than the expected density ($\delta \sim 10$) in the filaments, the result demonstrates that Suzaku can give a reasonable constraint on the soft excess emission.

5.3 Abell 2052

Suzaku carried out observations of Abell 2052 with 4 offset pointings to cover the cluster outskirts to about $20'$ (830 kpc at the source). Since the observations were all carried out only

40 days after the launch, the contamination in the XIS filter was not so much a problem. At the carbon K-edge energy, the reduction of the transmission efficiency was less than 8% (see Koyama et al. 2007).

The data were analysed by Tamura et al. (2007). By assuming a single temperature for the ICM and two temperatures (0.2 and 0.6 keV) for the Galactic component, they found a significant soft excess, with a spectrum well described by a featureless continuum modelled (for convenience reasons only) by pure thermal Bremsstrahlung with a temperature around 0.7 keV. The intensity of this soft component is consistent with a constant value over the entire observed field and stronger than the Galactic emission below 0.5 keV. If this emission is associated with Abell 2052, then its luminosity can be comparable to the ICM emission which shows $L_X \sim 1.4 \times 10^{44}$ erg s^{-1} in the 2–10 keV band. The cluster is located fairly close to the extension of the North Polar Spur (NPS), a large spur-like region with strong soft X-ray emission (see e.g. Willingale et al. 2003). However, the Suzaku spectrum of the NPS region showed the temperature to be about 0.3 keV with a fairly strong Mg-K emission line. This is inconsistent with the Abell 2052 soft excess spectrum.

A strong soft component without emission lines may be caused by extended non-thermal emission. Since the cluster shows neither strong radio emission nor merger features, such a non-thermal emission could be due to a rather old population of non-thermal electrons (see e.g. Rephaeli et al. 2008—Chap. 5, this issue).

5.4 Other Clusters

The Sculptor supercluster was observed with Suzaku in 4 pointings with the fields partially overlapping and connecting 3 main clusters (Abell 2811 at $z = 0.1086$, Abell 2804 at $z = 0.11245$ and Abell 2801 at $z = 0.11259$). The data were analysed by Kelley et al. (2007). The combined spectrum after removing bright clusters and point sources shows an excess in the energy range 0.6–1 keV, with a temperature of about 0.8 keV. Since the observed volume is extremely large ($\sim$500 Mpc^3), the uniform electron density implied is as low as 8×10^{-6} cm^{-3}.

The Suzaku XIS data on Abell 1060 ($z = 0.01140$, Sato et al. 2007), AWM7 ($z = 0.01724$, Sato et al. 2008), and Fornax (Matsushita et al. 2007) have been analysed in some detail. Even though the spectrum and intensity of the foreground Galactic emission have fairly large ambiguities, the observed energy spectra for these clusters are generally well fit by thermal models at energies down to about 0.3 keV. For the central region of Fornax ($r < 2'$), a two temperature model with k$T = 1.5$ and 0.8 keV was preferred. However, excess emission which causes significant deviation from these standard spectral fits has not been detected from these clusters.

Fujita et al. (2007) observed the region between Abell 399 and Abell 401 with Suzaku. They found no evidence for oxygen emission from the WHIM in this region and obtained a strict upper limit of 4.1×10^{-5} cm^{-3} on its density.

6 Soft X-Ray Emission Based on Chandra Observations

Reports on soft excess emission in clusters as observed by Chandra are scarce. Henriksen et al. (2004) detected soft excess emission in the merging cluster Abell 754 ($z = 0.0535$). Although a non-thermal origin could not be fully excluded, from their combined Chandra, BeppoSAX, ASCA and ROSAT PSPC analysis they conclude that a weak soft thermal component ($\sim$1 keV) is present in this hot ($\sim$10 keV) cluster. The emission extends out to 8′

from the core and is peaked in the cluster centre. Henriksen et al. attribute the emission to embedded groups of galaxies.

In another cluster, Abell 2163 ($z = 0.2030$), Henriksen and Hudson (2004) also report a soft and hard X-ray excess in their combined Chandra and ROSAT PSPC analysis. The excess can be modelled by a non-thermal component with photon index 2.7–5.9. Little detail about the observations is given in this paper, but the authors attribute the non-thermal emission to a weak merger shock that is present in this cluster.

Bonamente et al. (2007) analysed Chandra data of Abell 3112 in the central 1.0–2.5 arcmin region and found significant soft excess, reaching 20% of the hot gas emission level at 0.3 keV. Its spectrum is acceptably fit with a low temperature (∼0.5 keV) thermal model or by a non-thermal model ($\alpha_{\rm ph} \sim 1.8$). The best-fit values differ from those obtained with XMM-Newton data, indicating that remaining calibration uncertainties affect fine details. Nevertheless, the magnitude of the soft excess is above the level of calibration inaccuracies, and similar to that in XMM-Newton, thus proving the celestial origin of the soft excess in Abell 3112.

7 Discussion

Soft excess emission, defined here as excess emission over the usual thermal emission from hot intracluster gas at energies below 1 keV, has been detected significantly in ∼30% of clusters of galaxies. This value is based on the two largest cluster soft excess samples (Bonamente et al. 2003; Kaastra et al. 2003). However, as described above, such detections are difficult because of instrumental issues and for many objects a controversy remains.

Thermal and non-thermal models have been proposed to account for this soft excess emission. We will briefly summarise them below using Coma and Sérsic 159-03 as examples.

7.1 Thermal Models

7.1.1 WHIM in Coma

The massive soft excess halo around Coma (Bonamente et al. 2003) is not well accounted for by a non-thermal model. Furthermore, Finoguenov et al. (2003) detected O VII and O VIII line emission in the outskirts of Coma. Thus Coma is a good case for a cluster where the soft excess is dominantly of thermal nature. Bonamente et al. (2003) proposed that the soft component may reside in external filamentary structures of warm-hot intergalactic medium (WHIM). These filaments are predicted by hydrodynamic simulations of formation and evolution of large-scale structures, extending over several megaparsecs, containing a large fraction of the current epoch's baryons (e.g. Cen and Ostriker 1999; Davé et al. 2001; Yoshikawa and Sasaki 2006). In WHIM simulations, the temperatures are in the range 10^5–10^7 K, consistent with the soft excess properties found in Coma. Assuming a filamentary geometry, Bonamente et al. (2003) derived densities of $n \sim 10^{-5}$–10^{-4} cm^{-3} for the warm gas in Coma. These values are consistent with those in the WHIM simulations (10^{-7}–10^{-4} cm^{-3}). Assuming $n = 10^{-4}$ cm^{-3}, the implied mass of the filament exceeds that of the hot gas by a factor of 3.

7.1.2 Soft Emission from Merging Sub-halos

In Sérsic 159-03, detailed WHIM filament calculations showed that unrealistically long structures (∼100–1000 Mpc), projected on the line of sight of the cluster, are required to explain the soft excess detected in XMM-Newton data (Bonamente et al. 2005a). The assumption that the warm gas occupies the same volume as the hot gas has the problem that there is a pressure difference between the two components, and that the cooling time is short. As proposed by these authors, this problem can be avoided if one assumes that the warm gas is not distributed evenly but rather in high density 10^{-2}–10^{-3} cm^{-3} clumps with a volume filling factor $f \ll 1$. Such a distribution is predicted by the simulations of Cheng et al. (2005), in which the soft excess emission comes from high-density and low entropy gas associated with merging sub-groups, which preserve their identity before being destroyed and thermalised in the hot ICM. The simulated 0.2–1.0 keV band soft excess in the radial range of 0–0.5$r_{\rm vir}$ is consistent with that published in Bonamente et al. (2005a) for Sérsic 159-03. This scenario yields a warm gas mass of 25% of the hot gas mass for Sérsic 159-03.

However, Werner et al. (2007) show that the soft excess emission peaks at the position of the central cD galaxy and does not show any significant azimuthal variations. Moreover the soft excess in Sérsic 159-03 is observed out to radii of at least 1 Mpc. If this soft excess is associated with the gas of an infalling group, then this group is moving exactly along the line of sight. However, such an infalling group cannot explain the presence of the soft excess emission at large radii. Therefore the soft excess observed in Sérsic 159-03 is most probably not of thermal origin.

7.2 Non-thermal Models

Inverse Compton models of energetic electrons on CMB photons and/or on galaxy starlight have been developed by various authors to account for the soft excess observed in several of the objects presented here.

Non-thermal inverse Compton emission has a power-law spectrum with a relative flux which in the 0.3–10.0 keV band may account for more than 30% of the cluster emission. The best-fit power-law photon indices of soft excess clusters are typically between $\alpha_{\rm ph} \sim 2.0$–2.5. In the IC model, this corresponds to a differential relativistic electron number distribution $\mathrm{d}N/\mathrm{d}E = N_0 E^{(-\mu)}$ with $\mu = 3$–4. These are steeper than the distribution of the Galactic cosmic-ray electrons ($\mu \sim 2.7$). The steeper power-law distribution might indicate that the relativistic electrons suffered radiative losses (Sarazin 1999). Relativistic electrons in this energy range have relatively long lifetimes of $t_{\rm IC} = 2.3 \times 10^9 (\gamma/10^3)^{-1}(1+z)^{-4}$ yr and $t_{\rm syn} = 2.4 \times 10^{10}(\gamma/10^3)^{-1}(B/(1\ \mu G))^{-2}$ yr for inverse-Compton and synchrotron processes respectively.

As mentioned earlier, Werner et al. (2007) conclude that a non-thermal model best explains the observed properties of the soft excess in Sérsic 159-03. The total energy in relativistic electrons needed to explain the excess emission within the radius of 600 kpc does not exceed 1×10^{61} erg, while the total thermal energy within the same radius is 3×10^{63} erg. This means that even if the energy in relativistic ions is as much as ∼30 times larger than that in relativistic electrons, the total energy in cosmic ray particles will only account for 10% of the thermal energy of the ICM.

Models that may account for the observed soft excess can also be found in the following list: Enßlin et al. (1999); Atoyan and Völk (2000); Sarazin and Kempner (2000); Takizawa and Naito (2000); Fujita and Sarazin (2001); Petrosian (2001); De Paolis et al. (2003); Bowyer et al. (2004); Petrosian et al. (2008)—Chap. 10, this issue.

7.3 Some Problems and Open Questions

In the Virgo cluster a strong soft excess was detected in the extreme ultraviolet with EUVE (Lieu et al. 1996a; Berghöfer et al. 2000a; Bonamente et al. 2001b; Durret et al. 2002) and in the 0.2–0.4 keV band with ROSAT (Bonamente et al. 2002). However, the observations with XMM-Newton did not confirm the existence of this soft excess: a thermal model for the hot cluster emission with Galactic absorption describes the soft band X-ray spectra (above 0.3 keV) of Virgo obtained with XMM-Newton sufficiently well (Kaastra et al. 2003; Matsushita et al. 2002).

Arabadjis and Bregman (1999) claimed that some of the X-ray absorption cross sections were wrong, and that soft excesses would disappear when using the proper values. However, it is surprising that this claim has neither been confirmed nor refuted ever since. During some time, there have been wrong He cross sections by Balucińska-Church and McCammon (1992) in the XSPEC software (see http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/). Arabadjis & Bregman refer to those wrong cross sections. They have been improved by Yan et al. (1998) and are now included properly for example in Wilms' cross sections that are in XSPEC. They are also properly included in the SPEX software. Note that Wilms' more recent cross sections agree quite well with the older Morrison and McCammon (1983) work (see the discussion in Wilms et al. 2000). Lesson to be learned: it is important to check for each paper which cross sections/absorption model have been used!

Several searches for far-ultraviolet emission lines expected from a 10^6 K gas were performed with the FUSE satellite on the cores of several clusters. Oegerle et al. (2001) have reported the detection of O VI $\lambda = 1032$ Å in Abell 2597 ($z = 0.0824$), implying a mass inflow rate of about 40 $M_\odot$/yr. However, FUSE has not detected warm gas in five other clusters: Abell 1795 (Oegerle et al. 2001), Coma and Virgo (Dixon et al. 2001), and Abell 2029 ($z = 0.0775$) and Abell 3112 (Lecavelier des Etangs et al. 2004); the upper limits on the inflow rates for these five clusters were of the order 25 $M_\odot$/yr.

A recent review by Bregman (2007) summarises current knowledge on the search for missing baryons at low redshift. Note however that in his Fig. 9 (already in Bregman and Lloyd-Davies 2006) he claims that the soft excess found in some clusters by Kaastra et al. (2003) is an artefact due to wrong background subtraction. The argument is that there is a correlation between the Rosat R12 (low energy) count rate and the presence or absence of a soft excess. However, a closer inspection shows that this correlation is driven by a few clusters with strong Galactic absorption. As Kaastra et al. (2003) pointed out, in addition to the atomic gas visible at 21 cm these clusters also have significant X-ray opacity contributions due to dust or molecules, which naturally explains the observed flux deficit. When these clusters are excluded, the correlation fades away.

8 Note Added in Proof

Until very recently, the Galactic neutral hydrogen absorption taken into account when fitting X-ray spectra was extracted from the Dickey and Lockman (1990, hereafter DL90) survey. A better all-sky survey, the Leiden/Argentine/Bonn (LAB) Galactic H I Survey, has now become available (Kalberla et al. 2005). Though its use is not yet worldwide, it has been pointed out that, at least in some cases, the LAB survey gives notably smaller values than DL90 for the H I absorption. For example, in the case of Abell 3112, the LAB value is 1.3×10^{20} cm^{-2} while the DL90 value is 2.6×10^{20} cm^{-2}, and for Sérsic 159-03 (= AS 1101), the LAB value is 1.14×10^{20} cm^{-2}, while the DL90 value is 1.79×10^{20} cm^{-2}. Such changes

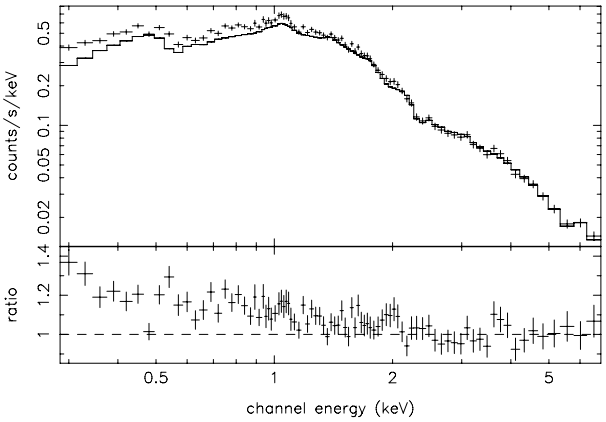


Fig. 5 XMM-Newton MOS data of Abell 3112 in a 1.5–2.9 arcmin annulus and fit with a single temperature MEKAL model in the 2–7 keV band with the LAB H I absorption column, showing a 20% soft excess

are expected to modify the soft excess emission derived from spectral fits. Fortunately, for many other clusters the differences between the H I absorption measured by LAB and by DL90 are negligible.

For Abell 3112, a fit of the XMM-Newton MOS data in a 1.5–2.9 arcmin annulus made by J. Nevalainen with a single temperature MEKAL model in the 2–7 keV band and extrapolated down to 0.3 keV using the LAB H I absorption value shows that a 20% soft excess remains (see Fig. 5). In order to eliminate the soft excess, the H I absorption should take the unrealistically low value of 1.5 10^{19} cm^{-2}.

For Sérsic 159-03, new fits of the joint Suzaku and XMM-Newton data in a 3–8 arcmin annulus were made by N. Werner using the new H I LAB absorption value. If the absorption is a free parameter in the fit, N_H goes to zero, and no good fit can be obtained with $N_H = 1.14 \times 10^{20}$ cm^{-2}. Adding a thermal component with a temperature of 0.2 keV improves the fit significantly (from a reduced χ^2 of 2.8 to 1.5), and the soft component is found to be present at a significance level of 13.5σ. If the fit is made with a multi-temperature model that accounts for the cluster emission with the absorption left as a free parameter, a 2σ upper limit for N_H of 1.4×10^{19} cm^{-2} is obtained; this value is inconsistent with the new LAB survey and clearly unphysically low. In the same way, if a power law is fit to the data, even if its total flux is reduced by about 20% it remains significant at a $\sim 10\sigma$ level. A soft component is therefore required even for zero absorption, implying that the soft excess is really strong in this cluster.

Therefore, even if they are somewhat reduced, the soft excesses observed in Abell 3112 and Sérsic 159-03 remain when taking into account the new values of the H I absorption.

9 Conclusions

While in most cases the origin of the soft excess emission is difficult to prove unambiguously, due to problems with instrument calibration and unknown background/foreground emission level, the independent detection of the phenomenon with many instruments gives confidence in the genuine nature of the phenomenon in a number of clusters. Both thermal and non-thermal emission mechanisms are probably at work in producing the soft excess emission in clusters.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". F.D. acknowledges support from CNES. SRON is supported financially by NWO, the Netherlands Foundation for Scientific Research. J.N. acknowledges support from the Academy of Finland. We thank R. Lieu, M. Bonamente and A. Fabian for pointing out the possible modifications of the soft excess emission when using more recent determinations of the hydrogen absorption along the line of sight (see note added in proof).

References

J.S. Arabadjis, J.N. Bregman, Astrophys. J. **514**, 607 (1999)
A.M. Atoyan, H.J. Völk, Astrophys. J. **535**, 45 (2000)
M. Balucińska-Church, D. McCammon, Astrophys. J. **400**, 699 (1992)
T.W. Berghöfer, S. Bowyer, Astrophys. J. **565**, L17 (2002)
T.W. Berghöfer, S. Bowyer, E. Korpela, Astrophys. J. **535**, 615 (2000a)
T.W. Berghöfer, S. Bowyer, E. Korpela, Astrophys. J. **545**, 695 (2000b)
M. Bonamente, R. Lieu, J.P.D. Mittaz, Astrophys. J. **546**, 805 (2001a)
M. Bonamente, R. Lieu, J.P.D. Mittaz, Astrophys. J. **547**, L7 (2001b)
M. Bonamente, R. Lieu, J. Nevalainen, J.S. Kaastra, Astrophys. J. **552**, L7 (2001c)
M. Bonamente, R. Lieu, J.P.D. Mittaz, Astrophys. J. **561**, L63 (2001d)
M. Bonamente, R. Lieu, M.K. Joy, J.H. Nevalainen, Astrophys. J. **576**, L688 (2002)
M. Bonamente, M.K. Joy, R. Lieu, Astrophys. J. **585**, 722 (2003)
M. Bonamente, R. Lieu, J. Mittaz, J.S. Kaastra, J. Nevalainen, Astrophys. J. **629**, 192 (2005a)
M. Bonamente, R. Lieu, J. Kaastra, Astron. Astrophys. **443**, 29 (2005b)
M. Bonamente, J. Nevalainen, R. Lieu, Astrophys. J. **668**, 796 (2007)
S. Bowyer, T.W. Berghöfer, Astrophys. J. **506**, 502 (1998)
S. Bowyer, R.F. Malina, in *Extreme Ultraviolet Astronomy*, ed. by R.F. Malina, S. Bowyer (Pergamon, New York, 1991), p. 397
S. Bowyer, A. Vikhlinin, J. Korean Astron. Soc. **37**, 579 (2004)
S. Bowyer, T.W. Berghöfer, E.J. Korpela, Astrophys. J. **526**, 592 (1999)
S. Bowyer, E.J. Korpela, T.W. Berghöfer, Astrophys. J. **548**, L35 (2001)
S. Bowyer, E.J. Korpela, M. Lampton, T.W. Jones, Astrophys. J. **605**, 168 (2004)
J.N. Bregman, Annu. Rev. Astron. Astrophys. **45**, 221 (2007)
J. Bregman, E.J. Lloyd-Davies, Astrophys. J. **644**, 167 (2006)
U. Briel, B. Aschenbach, G. Hasinger et al., http://heasarc.gsfc.nasa.gov/docs/rosat/ruh/handbook/rosathandbook.html (1996)
R. Cen, J.P. Ostriker, Astrophys. J. **514**, 1 (1999)
L.-M. Cheng, S. Borgani, P. Tozzi et al., Astron. Astrophys. **431**, 405 (2005)
R. Davé, R. Cen, J.P. Ostriker et al., Astrophys. J. **552**, 473 (2001)
F. De Paolis, G. Ingrosso, A.A. Nucita, D. Orlando, Astron. Astrophys. **398**, 435 (2003)
J. de Plaa, N. Werner, A.M. Bykov et al., Astron. Astrophys. **452**, 397 (2006)
J.M. Dickey, F.J. Lockman, Annu. Rev. Astron. Astrophys. **28**, 215 (1990)
W.V.D. Dixon, S. Sallmen, M. Hurwitz, R. Lieu, Astrophys. J. **550**, L25 (2001)
F. Durret, E. Slezak, R. Lieu, S. Dos Santos, M. Bonamente, Astron. Astrophys. **390**, 397 (2002)
M. Ehle, R. Breitfellner, R. González-Riestra et al., http://xmm.vilspa.esa.es/external/xmm_user_support/documentation/uhb/index.html (2006)
T.A. Enßlin, R. Lieu, P.L. Biermann, Astron. Astrophys. **344**, 409 (1999)
A. Finoguenov, U.G. Briel, J.P. Henry, Astron. Astrophys. **410**, 777 (2003)
A. Finoguenov, U.G. Briel, J.P. Henry et al., Astron. Astrophys. **419**, 47 (2004)
Y. Fujita, C.L. Sarazin, Astrophys. J. **563**, 660 (2001)
Y. Fujita, N. Tawa, K. Hayashida et al., Publ. Astron. Soc. Jpn. (2007 in press) (arXiv:0705.2017)
R. Fujimoto, K. Mitsuda, D. McCammon et al., Publ. Astron. Soc. Jpn. **59**, S133 (2007)
M.J. Henriksen, D.S. Hudson, J. Korean Astron. Soc. **37**, 299 (2004)
M.J. Henriksen, D.S. Hudson, E. Tittley, Astrophys. J. **610**, 762 (2004)
J.S. Kaastra, R. Lieu, J.P.D. Mittaz et al., Astrophys. J. **519**, L119 (1999)
J.S. Kaastra, C. Ferrigno, T. Tamura et al., Astron. Astrophys. **365**, L99 (2001)
J.S. Kaastra, R. Lieu, J.A.M. Bleeker, R. Mewe, S. Colafrancesco, Astrophys. J. **574**, L1 (2002)
J.S. Kaastra, R. Lieu, T. Tamura, F.B.S. Paerels, J. den Herder, Astron. Astrophys. **397**, 445 (2003)
J.S. Kaastra, R. Lieu, T. Tamura, F.B.S. Paerels, J.W.A. den Herder, in *Soft X-Ray Emission from Clusters of Galaxies and Related Phenomena*, ed. by R. Lieu, J. Mittaz, Astrophysics & Space Science Library, vol. 309 (2004), p. 37

J.S. Kaastra, F. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-4
P.M.W. Kalberla, W.B. Burton, D. Hartmann et al., Astron. Astrophys. **440**, 775 (2005)
R.L. Kelley, K. Mitsuda, C.A. Allen et al., Publ. Astron. Soc. Jpn. **59**, S7 (2007)
K. Koyama, H. Tsunemi, T. Dotani et al., Publ. Astron. Soc. Jpn. **59**, S23 (2007)
A. Lecavelier des Etangs, Gopal-Krishna, F. Durret, Astron. Astrophys. **421**, 503 (2004)
R. Lieu, J.P.D. Mittaz, S. Bowyer et al., Astrophys. J. **458**, L5 (1996a)
R. Lieu, J.P.D. Mittaz, S. Bowyer et al., Science **274**, 1335 (1996b)
R. Lieu, W.-H. Ip, W.I. Axford, M. Bonamente, Astrophys. J. **510**, L25 (1999a)
R. Lieu, M. Bonamente, J.P.D. Mittaz, Astrophys. J. **517**, L91 (1999b)
R. Lieu, M. Bonamente, J.P.D. Mittaz et al., Astrophys. J. **527**, L77 (1999c)
R. Lieu, M. Bonamente, J.P.D. Mittaz et al., Astron. Astrophys. **364**, 497 (2000)
K. Matsushita, E. Belsole, A. Finoguenov, H. Böhringer, Astron. Astrophys. **386**, 77 (2002)
K. Matsushita, Y. Fukazawa, J.P. Hughes et al., Publ. Astron. Soc. Jpn. **59**, S327 (2007)
K. Mitsuda, M. Bautz, H. Inoue et al., Publ. Astron. Soc. Jpn. **59**, S1 (2007)
J.P.D. Mittaz, R. Lieu, F.J. Lockman, Astrophys. J. **498**, L17 (1998)
R. Morrison, D. McCammon, Astrophys. J. **270**, 119 (1983)
J. Nevalainen, R. Lieu, M. Bonamente, D. Lumb, Astrophys. J. **584**, 716 (2003)
J. Nevalainen, M. Bonamente, J. Kaastra, Astrophys. J. **656**, 733 (2007)
W.R. Oegerle, L. Cowie, A. Davidsen et al., Astrophys. J. **560**, 187 (2001)
V. Petrosian, Astrophys. J. **557**, 560 (2001)
V. Petrosian, Y. Rephaeli, A.M. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9327-2
Y. Rephaeli, J. Nevalainen, T. Ohashi, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9314-7
C.L. Sarazin, Astrophys. J. **520**, 529 (1999)
C.L. Sarazin, J.C. Kempner, Astrophys. J. **533**, 73 (2000)
C.L. Sarazin, R. Lieu, Astrophys. J. **494**, L177 (1998)
K. Sato, N.Y. Yamasaki, M. Ishida et al., Publ. Astron. Soc. Jpn. **59**, 299 (2007)
K. Sato, K. Matsushita, Y. Ishisaki, N.Y. Yamasaki, M. Ishida, S. Sasaki, T. Ohashi, Publ. Astron. Soc. Jpn. (2008 in press). arXiv:0707.4342
P.J. Serlemitsos, Y. Soong, K.-W. Chan et al., Publ. Astron. Soc. Jpn. **59**, S9 (2007)
L. Strüder, U. Briel, K. Dennerl et al., Astron. Astrophys. **365**, L18 (2001)
Y. Takei, J.P. Henry, A. Finoguenov et al., Publ. Astron. Soc. Jpn. **59**, S339 (2007a)
Y. Takei, J.P. Henry, A. Finoguenov et al., Astrophys. J. **655**, 831 (2007b)
M. Takizawa, T. Naito, Astrophys. J. **535**, 586 (2000)
T. Tamura, Y. Takei, K. Mitsuda, T. Ohashi, K. Matsushita, the Suzaku A 2052 team, *Proc. "Tracing Cosmic Evolution with Clusters of Galaxies: Six Years Later"*, Sesto, Italy, 25–29 June 2007
M.J.L. Turner, A. Abbey, M. Arnaud et al., Astron. Astrophys. **365**, L27 (2001)
B.J. Wargelin, M. Markevitch, M. Juda et al., Astrophys. J. **607**, 596 (2004)
N. Werner, J.S. Kaastra, Y. Takei et al., Astron. Astrophys. **468**, 849 (2007)
R. Willingale, A.D.P. Hands, R.S. Warwick, S.L. Snowden, D.N. Burrows, Mon. Not. R. Astron. Soc. **343**, 995 (2003)
J. Wilms, A. Allen, R. McCray, Astrophys. J. **542**, 914 (2000)
M. Yan, H.R. Sadeghpour, A. Dalgarno, Astrophys. J. **496**, 1044 (1998)
K. Yoshikawa, S. Sasaki, Publ. Astron. Soc. Jpn. **58**, 641 (2006)

Chapter 5
Nonthermal Phenomena in Clusters of Galaxies

Y. Rephaeli · J. Nevalainen · T. Ohashi · A.M. Bykov

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9314-7

Abstract Recent observations of high energy (> 20 keV) X-ray emission in a few clusters of galaxies broaden our knowledge of physical phenomena in the intracluster space. This emission is likely to be nonthermal, probably resulting from Compton scattering of relativistic electrons by the cosmic microwave background (CMB) radiation. Direct evidence for the presence of relativistic electrons in some 50 clusters comes from measurements of extended radio emission in their central regions. We briefly review the main results from observations of extended regions of radio emission, and Faraday rotation measurements of background and cluster radio sources. The main focus of the review are searches for nonthermal X-ray emission conducted with past and currently operating satellites, which yielded appreciable evidence for nonthermal emission components in the spectra of a few clusters. This evidence is clearly not unequivocal, due to substantial observational and systematic uncertainties, in addition to virtually complete lack of spatial information. If indeed the emission has its origin in Compton scattering of relativistic electrons by the CMB, then the mean magnetic field strength and density of relativistic electrons in the cluster can be directly determined. Knowledge of these basic nonthermal quantities is valuable for the detailed description of processes in intracluster gas and for the origin of magnetic fields.

Y. Rephaeli (✉)
School of Physics & Astronomy, Tel Aviv University, Tel Aviv 69978, Israel
e-mail: yoelr@wise.tau.ac.il

Y. Rephaeli
Center for Astrophysics and Space Sciences, University of California, San Diego, La Jolla, CA 92093-0424, USA

J. Nevalainen
Observatory, University of Helsinki, P.O. Box 14, 00014 Helsinki, Finland

T. Ohashi
Department of Physics, Tokyo Metropolitan University, 1-1 Minami-Osawa, Tokyo 192-0397, Japan

A.M. Bykov
A.F. Ioffe Institute for Physics and Technology, 194021 St. Petersburg, Russia

J. Kaastra (ed.), *Clusters of Galaxies: Beyond the Thermal View*. DOI: 10.1007/978-0-387-78875-3_5

Keywords Clusters: general · X-ray emission

1 Introduction

Clusters of galaxies are the largest bound systems and the most important link to the large scale structure (LSS) of the Universe. The detailed properties of clusters, such as the distributions of their various mass constituents—dark matter, hot intracluster (IC) gas, and galaxies—dynamics, and thermal structure, are of much interest both intrinsically, and for the understanding of the formation and evolution of the LSS. Moreover, detailed *astrophysical* knowledge of clusters is essential for their use as *precise* cosmological probes to measure global parameters, such as H_0, $\Omega_{\rm M}$, & Ω_Λ, and parameters characterising the primordial density fluctuation field.

Recent observations of many clusters of galaxies with the *Chandra* and *XMM* satellites at energies $\epsilon \leq 10$ keV, have significantly advanced our knowledge of the morphology and thermal structure of hot IC gas, the source of the cluster thermal Bremsstrahlung emission. The improved determinations of the gas temperature, density, and metal abundances from these observations significantly improve estimates of such important quantities as the total cluster mass and its gaseous and baryonic mass fractions.

As has been the case in galaxies, in clusters too a more physically complete understanding of these systems necessitates knowledge also of non-thermal (NT) quantities and phenomena in the IC space. Observational evidence for the relevance of these phenomena in clusters comes mostly from measurements of extended regions of radio emission, and from Faraday rotation (FR) of the plane of polarisation of radio sources seen through (or inside) clusters. Since the observed radio emission is clearly synchrotron-produced, its level and spectrum yield direct information on IC relativistic electrons and magnetic fields. Information on cluster magnetic fields (separately from relativistic electron properties) is obtained also from FR measurements. Compton scattering of cosmic microwave background (CMB) photons by the radio-emitting relativistic electrons boosts photon energies to the X-and-γ regions (e.g., Rephaeli 1979). The search for cluster NT X-ray emission has begun long ago (Rephaeli et al. 1987), but first clear indications for emission at energies $\epsilon \geq 20$ keV came only after deep dedicated observations of a few clusters with the *RXTE* and *BeppoSAX* satellites (beginning with analyses of observations of the Coma cluster (Rephaeli et al. 1999; Fusco-Femiano et al. 1999). Radio and NT X-ray observations provide quantitative measures of very appreciable magnetic fields and relativistic electron densities in the observed clusters. These results open a new dimension in the study of clusters.

This is a review of cluster NT X-ray observations and their direct implications, including prospects for the detection of γ-ray emission. The literature on cluster NT phenomena is (perhaps somewhat surprisingly) quite extensive, including several reviews of radio emission (e.g., Govoni and Feretti 2004) and cluster magnetic fields (e.g., Carilli and Taylor 2002), and a review of the current status of radio observations by Ferrari et al. 2008—Chap. 6, this issue. In order to properly address the comparison between magnetic field values deduced from radio observations and jointly from NT X-ray and radio measurements, we include here a brief summary of cluster radio observations. Measurements of EUV emission in a few clusters, and claims that this emission is by energetic electrons, are reviewed by Durret et al. 2008—Chap. 4, this issue. NT radiation processes are reviewed by Petrosian et al. 2008—Chap. 10, this issue, and relevant aspects of particle acceleration mechanisms are reviewed by Petrosian and Bykov 2008—Chap. 11, this issue.

2 Radio Observations

Measurements of synchrotron radio emission at several frequencies usually provide the first evidence for the presence of a significant population of relativistic electrons and magnetic fields. This has been the prime evidence for IC fields and electrons. In addition, Faraday Rotation (FR) of the plane of polarisation of radiation from cluster and background radio galaxies has provided mostly statistical evidence for IC magnetic fields. Main results from these very different sets of observations are briefly reviewed in the next two subsections.

2.1 Extended IC Emission

In clusters the task of actually determining that the observed emission is from extended IC regions is quite challenging, since high resolution observations of the various discrete radio sources (that are mostly in the member galaxies) are required, in addition to the (lower resolution) observations of the extended low brightness emission. The truly extended emission is mapped upon subtraction of the respective contributions of these sources to the spectral flux. Observations of extended radio emission had begun some 50 years ago with measurements of Coma C in the central region of the Coma cluster by Seeger et al. (1957), followed by various other measurements at frequencies in the range 0.01–4.85 GHz, including the first detailed study at 408 MHz and 1407 MHz by Willson (1970).

Appreciable effort has been devoted to measure extended emission regions in clusters, which has resulted in the mapping of centrally located ('halo') and other ('relic') regions in about 50 clusters. Some 32 of these were found in a VLA survey (Giovannini et al. 1999; Giovannini and Feretti 2000) of 205 nearby clusters in the ACO catalogue, only about a dozen of these were previously known to have regions of extended radio emission. Primary interest (certainly to us here) is in the former sources, which constitute a truly cluster phenomenon. A central extended radio region (which is somewhat inappropriately referred to as 'halo') typically has a size of ∼1–2 Mpc, and a luminosity in the range 10^{40}–10^{42} erg s^{-1} (for $H_0 = 70$ km s^{-1} Mpc^{-1}) over the frequency band ∼0.04–5 GHz. With radio indices usually in the range ∼1–2, the emission is appreciably steeper than that of (most) radio galaxies.

The clusters in which radio halos and relics have already been found include evolving systems with a substantial degree of subclustering, as well as well-relaxed clusters that seem to have attained hydrostatic equilibrium. Halo morphologies are also quite varied, from roughly circular (projected) configuration to highly irregular shape, as can be seen in the contour maps produced by Giovannini et al. (1999, 2006), Giovannini and Feretti (2000). The spatial variety is reflected also in the wide distribution of values of the spectral index across the halo; e.g. see the maps in Giovannini et al. (1999).

Obviously, the magnetic field strength and relativistic electron density cannot both be determined from radio measurements alone, unless it is assumed that they are related so both can be determined from a single observable. It is commonly assumed that the total energy density of particles (mostly, protons and electrons) is equal to that in the magnetic field. The validity of this equipartition assumption is not obvious, especially in clusters where the particles and fields may have different origins and evolutionary histories. Moreover, the complex nature of the radio emission, and the expected significant variation of the magnetic field strength and relativistic electron density across a halo, imply that only rough estimates of these quantities can be obtained from measurements of the (spectral) flux integrated over the halo. When available, values of the halo mean equipartition field, $B_{\rm eq}$, are also listed in Table 1; generally, these substantially uncertain values are at the few μG level.

2.2 Faraday Rotation

IC magnetic fields can also be estimated by measuring the statistical depolarisation and Faraday Rotation of the plane of polarisation of radiation from background radio sources seen through clusters (e.g. Kim et al. 1991), and also from radio sources in the cluster. FR measurements sample the line of sight component of the randomly oriented (and spatially dependent) IC fields, weighted by the gas density, yielding a mean weighted value which we denote by $B_{\rm fr}$. This quantity was estimated by analysing the rotation measure (RM) distribution of individual radio sources in several clusters, including Cygnus A (Dreher et al. 1987), Hydra A (Taylor and Perley 1993; Vogt and Enßlin 2003), A 119 (Feretti et al. 1999; Dolag et al. 1999), A 400, and A 2634 (Vogt and Enßlin 2003). Analyses of FR measurements typically yield values of $B_{\rm fr}$ that are in the range of 1–10 μG.

Most FR studies are statistical, based on measurements of radio sources that are inside or behind clusters. An example is the work of Clarke et al. (2001), who determined the distribution of RM values with cluster-centric distance for 27 radio sources within or in the background (15 and 12, respectively) of 16 nearby clusters. Analysis of this distribution (including comparison with results for a control sample of radio sources seen outside the central regions of the clusters in the sample) yielded a mean field value of $\sim$5–10$(\ell/10\,{\rm kpc})^{-1/2}$ μG, where ℓ is a characteristic field spatial coherence (reversal) scale. In further work the sample of radio sources was significantly expanded (to about 70; Clarke 2004). In comparing different measures of the mean strength of IC fields it should be remembered that the selective sampling of locally enhanced fields in high gas density regions in cluster cores broadens the RM distribution, resulting in overestimation of the field mean strength.

Deduced values of $B_{\rm fr}$ yield substantially uncertain estimates of the mean field across a halo. The major inherent uncertainties stem from the need to separate the several contributions to the total RM (including that which is intrinsic to the radio source), the unknown tangled morphology of IC fields and their spatial variation across the cluster, as well as uncertainty in modelling the gas density profile. A discussion of these uncertainties, many aspects of which have already been assessed in some detail (e.g., Goldshmidt and Rephaeli 1993; Newman et al. 2002; Rudnick and Blundell 2003; Enßlin et al. 2003; Murgia et al. 2004), is out of the scope of our review.

3 Nonthermal X-ray Emission

As has been noted long ago, Compton scattering of the CMB by relativistic electrons boosts photon energies to the X-ray and γ-ray regions. Measurement of this radiation provides additional information that enables the determination of the electron density and mean magnetic field directly, *without* the need to invoke equipartition. The mean strength of the magnetic field which is deduced from the radio and NT X-ray fluxes, $B_{\rm rx}$, is essentially a volume average over the emitting region. The detection of IC radio and NT X-ray emission sets the stage for a more meaningful study of the origin of magnetic fields and cosmic rays in extragalactic environments.

The focus of this review are measurements of NT X-ray emission in large extended IC regions. NT emission is obviously predicted within dominant radio galaxies in the centres of clusters and in their radio halos. This emission has been sought in several galaxies including NGC 1275 in the Perseus cluster (e.g., Sanders and Fabian 2007) and M 87 in Virgo (Simionescu et al. 2007), but does not seem to have been detected. The mapping of the electron spectrum in these inner regions clearly provides crucial information on the initial source

of IC relativistic electrons and most likely also energetic protons, whose energy input and heating of the gas could possibly be important in the inner core of Perseus (Rephaeli and Silk 1995) and other clusters. NT X-ray emission in these environments is quite complex due to its multi-source (AGN, jet, binaries, and halo), temporally variable nature. A discussion of NT emission in these essentially galactic environments is out of the scope of our review.

The prospects for measuring cluster NT X-ray emission motivated detailed calculations of the predicted emission (Rephaeli 1977, 1979). In these calculations the relativistic electron spectrum was directly related to the measured radio spectrum and no attempt was made to model the spectra of possible energetic electron populations. Energetic protons, which are a major Galactic cosmic ray component, are also expected in the IC space, particularly so in the inner cores of clusters dominated by a radio galaxy. Their interactions with protons in the gas produce neutral and charged pions, whose decays yield γ-ray emission and secondary electrons (e.g., Dennison 1980; Dermer and Rephaeli 1988; Blasi and Colafrancesco 1999). Energetic protons also deposit energy and heat IC gas through their Coulomb interactions with electrons and protons in the gas (Rephaeli 1987; Rephaeli and Silk 1995).

With mean field values in the range of 0.1–1 μG, the energy range of electrons emitting at the observed radio frequencies is ∼1–100 GeV. Of course, electrons with energies outside this range are also expected, either as part of the same or a different population. In particular, NT Bremsstrahlung EUV and X-ray emission by (the more numerous) lower energy electrons would also be expected (Kaastra et al. 1998; Sarazin 1999). However, at energies below ∼200 MeV, the main energy loss is electronic excitations in the gas (Rephaeli 1979); this sets a stringent limit on the contribution to the NT X-ray emission by a low energy electron population (Rephaeli 2001; Petrosian 2001). We will therefore assume that NT X-ray emission is largely due to Compton scattering of radio-emitting electrons by the CMB.

The main characteristics of the predicted Compton-produced emission are: (a) X-ray to radio flux ratio is roughly equal to the ratio of the CMB energy density to the magnetic field energy density. (b) Power-law index is nearly equal to the radio index. (c) Matching X-ray and radio centroids, with the X-ray spatial profile generally shallower than that of the radio emission. These features are expected only if the Compton-produced emission is identified and separated from other contributions, the primary thermal emission as well as NT emission from relics and individual galaxies. Due to source confusion, as well as other systematic and observational uncertainties, spatial information is *crucially* needed in order to clearly identify the origin of NT X-ray emission.

We first briefly summarise the required sensitivity and level of spatial resolution for the detection of NT emission, and then review the results of the search for this emission with the *HEAO-1*, *CGRO*, *RXTE*, *BeppoSAX*, *Suzaku*, and *INTEGRAL*.

3.1 Instrumental Requirements

Cluster X-ray emission is primarily thermal up to ∼30–40 keV; therefore, clear identification of NT emission necessitates also precise measurement of the thermal emission in order to account for and separate it from the sought NT component. The required detector spectral range must therefore extend down to sufficiently low energies.

Using standard expressions for Compton and synchrotron emission from the same population of relativistic electrons, the predicted level of Compton-produced emission from nearby clusters with measured radio halos can be readily estimated. Since imaging information at the high energy X-ray regime ($\epsilon > 30$) is minimal at best, we will ignore spatial

factors in the Compton-synchrotron formulae (e.g., Rephaeli 1979) that include integrals over the relativistic electron and magnetic fields spatial profiles. This is valid as long as the NT X-ray emission region coincides with the radio region. We base our estimate of the required detector sensitivity on the level of radio emission as measured in the nearby clusters with well measured halos (such as Coma, A 2256, A 2319), for which the feasibility of detection of the NT emission is likely to be highest. Using the range of values for radio spectral index and flux, we estimate the level of NT flux at 40 keV to be typically $\sim 1 \times 10^{-6}$ cm^{-2} s^{-1} keV^{-1} and an integrated flux of $\sim 3 \times 10^{-5}$ cm^{-2} s^{-1} in the 40–80 keV band, if the mean field (across the emitting region) is $B \sim 0.3$ μG. The value of the field is critical due to the strong B dependence of the flux, with a power-law exponent typically larger than 2. For example, the predicted flux in the Coma cluster is some 16 times lower if $B = 1$ μG.

Given the great difficulty in detecting the predicted NT flux at energies $\epsilon \geq 40$ keV, it is only reasonable to search for it also at lower energies. To do so optimally would require a high-sensitivity detector that covers a wider spectral band than that of any previously flown instrument. Combining measurements made with two different detectors is problematic, given the inherent difficulty in the requisite precise cross-calibration (at the ∼1% level) of signals from detectors on the same satellite, and even more so when the detectors were on different satellites altogether. An added difficulty arises in the extremely difficult task of detecting a weak secondary component at energies below ∼1 keV where photoelectric absorption is strong. This should be kept in mind when assessing results form such combined analyses, some of which are discussed below.

Even if NT emission is detected its origin has to be verified before it is identified with electrons in the halo. The first requirement is that the measured emission does not show any temporal variation (that would usually imply AGN origin). Sizes of radio halos in nearby clusters are in the range of 15′–30′, which in principle can be resolved by the *IBIS* instrument aboard *INTEGRAL*, the first imaging experiment in the high energy X-ray regime. Clearly, higher detector sensitivity is required for resolved measurements of such large regions. All the other presently active (and previous) satellites with high energy spectral capability have FOVs that are larger than the halos of (even) nearby clusters. Measurements of NT emission have so far been solely spectral, thus providing only the necessary—but not sufficient—evidence for the detection of NT emission in a few clusters.

3.2 Initial Search for NT X-ray Emission

Systematic searches for NT X-ray emission began with the analysis of archival *HEAO-1* measurements of six clusters (Coma, A 401, A 2142, A 2255, A 2256, A 2319) with measured radio halos (Rephaeli et al. 1987; Rephaeli and Gruber 1988), and continued with *CGRO* observation of the Coma cluster (Rephaeli et al. 1994). No significant NT emission was detected, resulting in lower limits on the mean, volume-averaged magnetic fields in these clusters, $B_{\rm rx} = O(0.1\ \mu{\rm G})$.

Observation of the Coma cluster with the *CGRO/OSSE* experiment for ∼380 ks was the first dedicated measurement of a cluster aimed at measuring NT emission. The fact that the predicted emission was not detected clearly established that instruments with higher detector sensitivity, significantly lower level of internal background, and much smaller FOV (than that of *OSSE*, $\sim 1^\circ \times 1^\circ$ FWHM), are the minimal requirements for measurement of emission in the 40–100 keV band.

Attempts to measure NT emission from clusters continued with all subsequent X-ray satellites whose nominal spectral band extended beyond 10 keV. First observational evidence

for NT emission came from long measurements with the *RXTE* and *BeppoSAX* satellites, as discussed in the next two subsections, followed by a review of results from *ASCA*, initial results from *Suzaku/HXD*, and the first spatial analysis of *INTEGRAL/IBIS* observations of the Coma cluster.

3.3 Search for NT Emission with *RXTE*

The search for cluster NT emission advanced significantly through long observations with the PCA and HEXTE experiments aboard the *RXTE* satellite. These instruments have the minimally required capabilities for detection of a weak NT spectral component—sufficiently high sensitivity for precise measurement of the primary thermal emission, and good background subtraction achieved by (on and off source) 'rocking' of the HEXTE detectors. The crucial *RXTE* features which are essential for identifying the small NT component are spectral overlap of the two experiments over the effective (narrower than nominal) ∼13–25 keV band, and the same triangular spatial response function with 58′ FWHM. As has been noted already, the *RXTE* does not have the capability to resolve the emission even in nearby (rich) clusters, implying that the exact origin of NT emission cannot be identified even when a secondary spectral component is clearly detected.

We review here results of searches for NT components in the spectra of clusters that were observed for at least ∼100 ks. Shorter observations were mostly aimed at measurement of the primary thermal emission, with the exception of a 70 ks observation of A 754 (Valinia et al. 1999), and even shorter ∼30 ks observations of A 2256 (Henriksen 1999) and A 1367 (Henriksen and Mushotzky 2001), which resulted in upper limits on NT emission in these clusters. For obvious reasons, Coma was the first cluster searched for NT emission with *RXTE* (Rephaeli et al. 1999). The search continued with long observations of A 2319, A 2256, and A 2163, all with observed radio halos. Results of the search for NT emission in these and other clusters are listed in Table 1, and briefly described below.

Coma The cluster was initially observed (in 1996) for ∼90 ks with the PCA, and for ∼29 ks with HEXTE. Analysis of these measurements showed evidence for the presence of a second spectral component at energies up to ∼20 keV (Rephaeli et al. 1999), but the nature of the secondary emission could not be determined. In order to improve the quality

Table 1 NT emission parameters from *RXTE* measurements and (deduced mean radio-and-X-determined field) $B_{\rm rx}$

Cluster	20–80 keV flux (10^{-12} erg s^{-1} cm^{-2})	Γ	$B_{\rm rx}$ (μG)	Reference
Coma	21 ± 6	2.1 ± 0.5	0.1–0.3	Rephaeli and Gruber (2002)
A 2163	11^{+17}_{-9}	$1.8^{+0.9}_{-4.2}$	0.4 ± 0.2	Rephaeli et al. (2006)
A 2256	4.6 ± 2.4	$2.2^{+0.9}_{-0.3}$	$0.2^{+1.0}_{-0.1}$	Rephaeli and Gruber (2003)
A 2319	14 ± 3	2.4 ± 0.3	0.1–0.3	Gruber and Rephaeli (2002)
A 3667	≤ 4	∼2.1	≥ 0.4	Rephaeli and Gruber (2004)
1ES 0657-55.8	$\sim 5 \pm 3$	1.6 ± 0.3	1.2[a]	Petrosian et al. (2006)

All quoted errors are at the 90% confidence level

[a]The magnetic field value in 1ES 0657-558 was derived assuming energy equipartition between the field and particles

of the spectral results, the cluster was observed again (in 2000) for ∼177 ks. Analysis of the new observations (Rephaeli and Gruber 2002, hereafter RG02) yielded results that are consistent with those from the first observation. Joint analysis of the full *RXTE* dataset shows that—from a statistical point of view—no preference could be determined for the nature of the secondary component. If thermal, the most probable temperature would be very high, $kT_2 \simeq 37.1$ keV, but at its lowest boundary, the 90% contour region includes a much lower, physically more acceptable value. However, the best fit combination of the temperatures of the primary and secondary components is then $kT_1 = 5.5$ keV, and $kT_2 = 9$ keV, respectively, with the respective 4–20 keV fractional fluxes of ∼24% and ∼76%. RG02 argued that it is quite unlikely that about a quarter of the emission comes from gas at a significantly lower temperature than the mean value deduced by virtually all previous X-ray satellites. In particular, such a component would have been detected in the high spatially resolved *XMM* and *Chandra* maps of the cluster.

To quantitatively assess the possibility that a two-temperature gas model is just a simplified representation of a more realistic continuous temperature distribution, RG02 have assumed a polytropic gas temperature profile of the form $T(r) \propto n(r)^{\gamma-1}$, with a β profile for the gas density, $n(r) \propto (1 + r^2/r_c^2)^{-3\beta/2}$, where r_c is the core radius. With these profiles they calculated the integrated flux and the mean emissivity-weighted temperatures as functions of γ, β, and r. These quantities were then calculated in the projected (2D) circular regions $[0, R]$ and $[R, R_0]$ by convolving over the triangular response of the PCA with $R_0 \simeq 58'$. From ROSAT observations, $r_c \sim 10.0'$, and $\beta \simeq 0.70 \pm 0.05$ (Mohr et al. 1999). The range of values of R, β, and γ for which the two mean emissivity-weighted temperatures and respective fluxes from these regions were closest to the values deduced from the spectral analysis, were then determined. The results of these calculations indicate that for $0.5 \leq \beta \leq 0.9$ and $1 \leq \gamma \leq 5/3$, there is no acceptable polytropic configuration that matches the observationally deduced values of the temperatures and fractional fluxes. For low values of γ the temperature gradient is too shallow, whereas for high γ values the implied central temperature is unrealistically high. Based on these results RG02 concluded that the two thermal components model does not seem to be consistent with the *RXTE* measurements, but certainly could not be ruled out.

Clearly, a search for NT emission motivated the long observation, and this alternative origin was assessed in detail by RG02. Since emission from an AGN in the FOV was considered unlikely, they assessed whether the secondary emission is due to Compton scattering of relativistic electrons whose presence in Coma is directly inferred from many measurements of spatially extended region of radio emission (e.g. Kim et al. 1991; Giovannini et al. 1993; Thierbach et al. 2003). From the measured radio spectral index, 1.34 ± 0.1, it follows that the predicted power-law (photon) flux from Compton scattering of these electrons by the CMB should have an index $\Gamma \sim 2.3$. A power-law fit to the secondary component yielded 2.1 ± 0.5 at the 90% confidence level (CL). Results of this fit are shown in Fig. 1.

With the measured mean radio flux of 0.72 ± 0.21 Jy at 1 GHz, and the deduced level of power-law X-ray flux, the mean volume-averaged value of the magnetic field, B_{rx}, was computed to be in the range 0.1–0.3 μG. As noted by RG02, this estimate is based on the assumption that the spatial factors in the theoretical expressions for the two fluxes (Rephaeli 1979) are roughly equal, an implicit assumption made in virtually all previous attempts to derive the strength of cluster fields from radio and X-ray measurements.

A similar procedure was employed in analyses of observations of the other clusters (A 2319, A 2256, and A 2163) that were initiated by Rephaeli & Gruber.

A 2319 The cluster was observed (in 1999) for ∼160 ks. Analysis of the data (Gruber and Rephaeli 2002) showed no noticeable variability over the ∼8 week observation. The

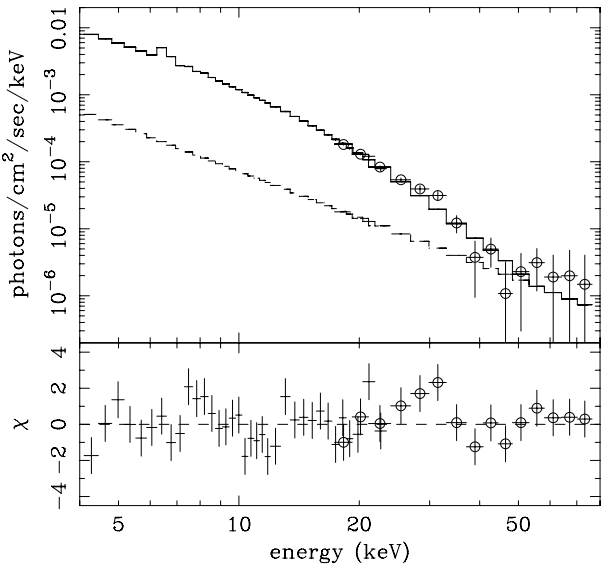


Fig. 1 The RXTE spectrum of the Coma cluster and folded Raymond-Smith (k$T \simeq 7.67$), and power-law (photon index $= 2.3$) models (from Rephaeli and Gruber 2002). HEXTE data points are marked with circles and 68% error bars. The total fitted spectrum is shown with a histogram, while the lower histogram shows the power-law portion of the best fit. The quality of the fit is demonstrated in the lower panel, which displays the observed difference normalised to the standard error of the data point

quality of the data allowed a meaningful search for emission whose spectral properties are distinct from those of the primary thermal emission with measured mean temperature in the range 8–10 keV. A fit to a single thermal component yielded k$T = 8.6 \pm 0.1$ (90% CL), a low iron abundance $Z_{\rm Fe} \sim 0.16 \pm 0.02$, and large positive residuals below 6 keV and between 15 to 30 keV. The quality of the fit was very significantly improved when a second component was added. A two-temperature model yielded k$T_1 \simeq 10.1 \pm 0.6$, k$T_2 \simeq 2.8 \pm 0.6$, and $Z_{\rm Fe} \sim 0.23 \pm 0.03$, a value consistent with previously measured values. An equally good fit was obtained by a combination of primary thermal and secondary NT components, with k$T \simeq 8.9 \pm 0.6$, and photon index $\Gamma \simeq 2.4 \pm 0.3$. Similar results were obtained when a joint analysis was performed of the *RXTE* data and archival ASCA data. The deduced value of Γ is consistent with the measured spectrum of extended radio emission in A2319. Identification of the power-law emission as Compton scattering of the radio-emitting electrons by the CMB resulted in $B_{\rm rx} \sim 0.1$–0.3 μG, and $\sim 4 \times 10^{-14} (R/2\ {\rm Mpc})^{-3}$ erg cm^{-3} for the mean energy density of the emitting electrons in the central region (radius R) of the cluster.

A 2256 Following a very short ∼30 ks observation of A 2256 (in 1997)—which resulted in an upper limit on NT emission (Henriksen 1999)—the cluster was observed (in 2001 and 2002) for a total of ∼343 ks (*PCA*) and ∼88 ks (*HEXTE*). The data analysis (Rephaeli and Gruber 2003) yielded evidence for two components in the spectrum. Based on statistical likelihood alone the secondary component can be either thermal or power-law. Joint analysis of the *RXTE* and archival *ASCA* data sets yielded k$T_1 = 7.9^{+0.5}_{-0.2}$ and k$T_2 = 1.5^{+1.0}_{-0.4}$, when the second component is also thermal, and k$T = 7.7^{+0.3}_{-0.4}$ and $\Gamma = 2.2^{+0.9}_{-0.3}$ (90% CL), if the second component is a power-law. Identifying the secondary emission as due to Compton scattering of the radio producing relativistic electrons yielded $B_{\rm rx} \simeq 0.2^{+1.0}_{-0.1}$ μG in the central 1° region of the cluster, a region which contains both the halo and relic sources.

A 2163 The moderately-distant ($z = 0.203$) cluster was observed for ∼530 ks (during a 6 month period in 2004). Primary thermal emission in this cluster comes from very hot IC gas with k$T \sim 15$ keV, but analysis of the observations (Rephaeli et al. 2006) indicated very

clearly that this component does not by itself provide the best fitting model. A secondary emission component was needed, and while this could also be thermal at a temperature significantly lower than 15 keV, the best fit to the full dataset was obtained with a power law secondary spectral component. The parameters of the NT emission imply a significant fractional flux amounting to ∼25% of the integrated 3–50 keV emission. NT emission is expected given the intense level of radio emission, most prominently from a large radio halo. Rephaeli et al. (2006) assumed that the NT emission originates in Compton scattering of (the radio-emitting) relativistic electrons by the CMB, and estimated $B_{\rm rx} \sim 0.4 \pm 0.2$ μG to be an overall mean field strength in the large complex region of radio emission in the cluster.

A 3667 The cluster was observed (in 2001 and 2002) for ∼141 ks; analysis of the RXTE observation and lower energy *ASCA* data, yielded only marginal evidence for a secondary power-law emission component in the spectrum (Rephaeli and Gruber 2004). This resulted in an upper limit of 2.6×10^{-12} erg cm^{-2} s^{-1} (at 90% CL) on NT emission in the 15–35 keV band. When combined with the measured radio flux and spectral index of the dominant region of extended radio emission, this limit implies a lower limit of ∼0.4 μG on the mean, volume-averaged magnetic field in A 3667.

1ES 0657-55.8 At $z = 0.296$ the 'bullet' cluster is the most distant cluster searched for NT emission with *RXTE*. The cluster was observed (in 2002 and 2003) for a total of ∼400 ks. Joint analysis of the *RXTE* observations together with archival *XMM-Newton* observations clearly indicated the presence of a second spectral component (Petrosian et al. 2006). While the nature of this second component cannot be determined from the spectral analysis alone, the authors argue that a power-law spectral shape is a more viable interpretation than an exponential. Since no radio data were available, the authors could only determine a mean field value of ∼1.2 μG assuming energy equipartition between the field and relativistic electrons.

3.4 Search for NT Emission with *BeppoSAX*

For observations of clusters the two relevant instruments were the MECS and PDS, the former with a spectral range of 1.3–10 keV, 56′ FOV and ∼2′ resolution, and the latter with response in the nominal 15–300 keV range and 1.3° FOV. The capability to separate the respective contributions of low and high energy components requires simultaneous and self-consistent measurements. For this reason the lack of spectral overlap of the two instruments (and their somewhat different FOV) was unfortunate.

3.4.1 Observations of Individual Clusters

Results of searches for cluster hard X-ray emission are summarised in Table 2; a brief discussion of these results follows.

Coma Fusco-Femiano et al. (1999) reported a 4.5σ detection of excess emission above the primary thermal emission in ∼91 ks PDS observations carried out in 1997, but a later re-analysis resulted in a reduced estimate of the significance to 3.4σ (Fusco-Femiano et al. 2004). The cluster was observed again for ∼300 ks (in 2000), and an analysis of the combined datasets—with a thermal component whose temperature was *assumed* to be $kT = 8.1$ keV—yielded a 4.8σ detection of excess emission in the 20–80 keV band (Fusco-Femiano et al. 2004). Assuming a power-law form for this excess emission with index

Table 2 NT emission parameters from *BeppoSAX* measurements and (deduced mean radio-and-X-determined field) $B_{\rm rx}$

Cluster	20–80 keV flux (10^{-12} erg s^{-1} cm^{-2})	Γ	$B_{\rm rx}$ (μG)	Reference
Coma[a]	15 ± 5	?	0.2	Fusco-Femiano et al. (2004)
A 2256	$8.9^{+4.0}_{-3.6}$	$1.5^{+0.3}_{-1.2}$	0.05	Fusco-Femiano et al. (2000)
A 2199	9.8 ± 4.0	1.8 ± 0.4	?	Kaastra et al. (1999)
A 2319	≤ 23	?	≥ 0.04	Molendi et al. (1999)
A 3667[b]	≤ 6.4	2.1	≥ 0.4	Fusco-Femiano et al. (2001)
A 754[c]	~ 2	?	0.1	Fusco-Femiano et al. (2003)
Centaurus[d]	10.2 ± 4.6	$1.5^{+2.3}_{-1.3}$		Molendi et al. (2002)

All quoted errors are at the 90% confidence level

[a]According to Rossetti and Molendi (2007) proper accounting for uncertainties makes the detection insignificant

[b]The emission from AGN FRL339 was not subtracted out

[c]The measured flux is possibly from the radio galaxy 26W20

[d]The measured flux is possibly due to an AGN

$\Gamma \sim 2.0$, the 20–80 keV flux of $1.5 \pm 0.5 \times 10^{-11}$ erg s^{-1} cm^{-2} was derived (a value which is consistent with that derived by Nevalainen et al. (2004). Interpreting this to be of Compton origin, and using the radio flux, resulted in $B_{\rm rx} \sim 0.2$ μG.

Rossetti and Molendi (2004) questioned the reliability of the Fusco-Femiano et al. (1999, 2004) analyses. Having carried out their own analysis of 69 blank fields observed with PDS, they found that the two offset positions used for the local background determination yielded systematically different results for the background-subtracted flux. Since Fusco-Femiano et al. (1999, 2004) used only one offset pointing (due to contamination in the other offset field), without any adjustment for the systematic difference, Rossetti and Molendi (2004) claimed that the results of Fusco-Femiano et al. are incorrect. Based on the use of a different software package (SAXDAS) for background determination, Rossetti and Molendi (2004) determined that the significance of detection of excess emission in the original observation was only at the 2σ level, while their analysis of the second observation yielded no evidence for excess emission, in strong contrast with the 4.8σ detection level deduced by Fusco-Femiano et al. (2004).

The analysis was repeated (see Fig. 2) by Fusco-Femiano et al. (2007); they concluded that the discrepant results are due to the presence of two variable sources in one of the background fields, which is why they ignored that field in their analysis. The different background fields and the use of the XAS software (rather than SAXDAS) were suggested to be the reasons for the high significance (4.8σ) detection reported by Fusco-Femiano et al. (2004). In a recent reply to Fusco-Femiano et al. (2007), Rossetti and Molendi (2007) defended their re-analysis and the lower detection significance of the hard excess, re-emphasising that the significance is highly sensitive to both the choice of the offset pointings for local background estimation, and also to the exact value of the temperature adopted in the analysis. Landi (2005) analysed a sample of 868 blank fields and found no difference in flux values when using either of the two offset pointings. However, Rossetti and Molendi (2007) could not investigate the reasons for the discrepancy between their blank sky measurements and those of Landi (2005) since the details of the latter work are not yet publicly available.

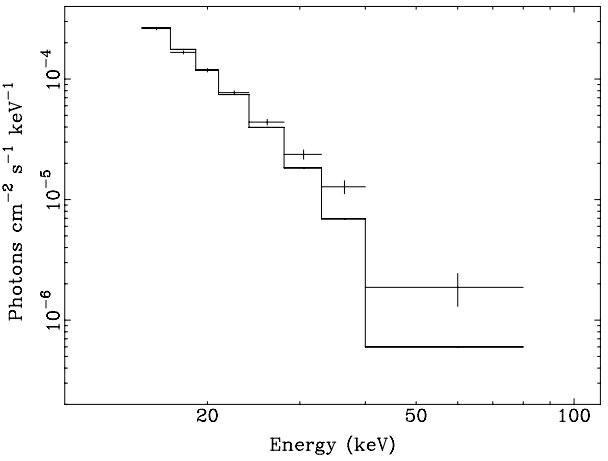


Fig. 2 *BeppoSAX* spectrum of the Coma cluster obtained using the XAS software package (from Fusco-Femiano et al. 2007). Data are from PDS measurements (above 15 keV, with 68% error bars); the *line* is a best-fit thermal model with $kT = 8.11$ keV obtained previously from *Ginga* measurements (David et al. 1993)

Flux confusion due to unidentified (and therefore unremoved) background AGN is particularly troublesome due to the relatively large FOV (1.3°) of the PDS. Although some of the integrated emission of distant AGN in the cluster field is removed in the rocking mode, there still is a residual signal due to AGN background fluctuations, and possibly also the presence of AGN in the cluster. This introduces a level of systematic uncertainty that was not specifically accounted for in most published PDS cluster analyses. Analysing 164 PDS blank-sky observations, Nevalainen et al. (2004) found a systematic difference in the background-subtracted fluxes when using only one of the two pointings available for the local background estimate in the standard rocking mode. The authors speculate that this might be due to either radiation entering the collimators from the side, screening of the instruments by the satellite, or the fact that the detector is looking at more/less radioactive parts of the satellite. These results are qualitatively similar to those of Rossetti and Molendi (2004), but are more pronounced by a factor of ∼2. The difference in flux values when using either of the two offset pointings seems to be the reason for the different results obtained for the significance of the detection of NT emission in Coma. The issue remains unsettled and needs to be further explored.

Moreover, as we have pointed out, precise separation between the primary and secondary spectral components depends very much on the ability to do a simultaneous fit to the parameters of both thermal and possibly NT emission. The search for a NT component with *BeppoSAX* suffered from lack of spectral overlap between the MECS and PDS instruments, and consequently the need to *adopt* a previously determined value of kT. Due to inherent differences between the *HEXTE* and *PDS* instruments, only a rough comparison between the respective results can be made. In this spirit it is apparent that the two estimates for the NT (20–80 keV) flux (Rephaeli and Gruber 2002; Fusco-Femiano et al. 2004) are in agreement (see Tables 1 and 2).

A 2256 Fusco-Femiano et al. (2005) analysed two separate PDS observations of the cluster for a total of ∼430 ks; they claimed a 4.8σ detection of excess emission. Fitting this excess to a power-law yielded $\Gamma = 1.5^{+0.3}_{-1.2}$ and a 20–80 keV flux of 8.9×10^{-12} erg cm^{-2} s^{-1}, which is essentially the same value deduced by Nevalainen et al. (2004). Use of the measured radio flux yielded $B_{\rm rx} \sim 0.05$ μG, assuming that the NT emission they deduced originates in the radio relic located in the NW side of the cluster. We note that the flux deduced by Rephaeli and Gruber (2003) is about a factor of ∼2 lower than that of Fusco-Femiano et al. (2005),

but given the large uncertainty in Γ, for which the former authors quote a relatively low value, this difference is not very significant.

A 2199 Analysis of MECS 8–10 keV measurements led Kaastra et al. (1999) to claim detection of a significant excess emission (over that of the best-fit thermal model). Fit of a power-law to the data yielded an index of 1.81 ± 0.25. They concluded that this emission originates in an outer region (beyond a 300 kpc central radial region), and that its relative strength (with respect to thermal emission) increases such that it dominates the 8–10 keV emission at ∼1 Mpc from the centre.

A 3667 A marginal detection (at a 2.6σ significance) of excess 20–35 keV emission was reported by Fusco-Femiano et al. (2001). A power-law fit resulted in an index of ∼2.1, from which an upper limit on NT 20–80 keV flux of 6.4×10^{-12} erg s^{-1} cm^{-2}, and a lower limit $B_{\rm rx} \simeq 0.4$ μG were deduced. However, the authors did not consider the impact of the presence of the Seyfert 1 galaxy FRL 339 in the PDS FOV; emission from this galaxy is at a level comparable to that attributed to the NT component (Nevalainen et al. 2004). This would seem indeed the case, given the fact that analysis of *RXTE* measurements yielded only an upper limit on NT emission (Rephaeli and Gruber 2004).

A 754 Fusco-Femiano et al. (2003) reported the detection of excess emission above 45 keV at a 3.2σ significance. They deduced a 10–40 keV flux of 2×10^{-12} erg s^{-1} cm^{-2} and $B_{\rm rx} \simeq 0.1$ μG, but noted that the presumed NT emission could possibly be from the radio galaxy 26W20.

Centaurus Molendi et al. (2002) detected a hard X-ray excess at the 3.6σ level. The best-fit power-law model yielded an index of $\Gamma = 1.5^{+1.4}_{-0.8}$, and a 20–200 keV flux of 2.2×10^{-11} erg s^{-1} cm^{-2}, but they concluded that the emission may originate in a serendipitous AGN.

A 2319 No statistically significant power-law emission was detected in the analysis of a very short ∼20 ks (PDS) observation, resulting in an upper limit of 2×10^{-11} erg s^{-1} cm^{-2}, and a lower limit $B_{\rm rx} \sim 0.04$ μG (Molendi et al. 1999). As discussed above, the much longer (∼160 ks) *RXTE* observation led to a significant NT flux.

3.4.2 Statistical Results from a Cluster Sample

An attempt to obtain some insight from co-added PDS data on a sample of 27 clusters was made by Nevalainen et al. (2004). A brief review of their results follows.

The problematic aspects of a statistical study of PDS data should first be summarised. As we have already noted, flux confusion due to unidentified AGN is of particular concern due to the relatively large PDS FOV. Nevalainen et al. (2004) used optical catalogues to identify Seyfert 1 galaxies, and BeppoSAX MECS or ROSAT PSPC data at their locations to estimate the AGN contribution in the PDS spectra. Clusters for which the estimated Seyfert 1 contribution was more than 15% of the total signal were removed from the initial sample. Population synthesis modelling of the cosmic X-ray background (CXB) indicates that 80% of the AGN need to be obscured to produce the CXB spectrum (Gilli et al. 1999), which is harder than the spectrum of unobscured AGN. Indeed, recent deep X-ray observations of blank fields (e.g., Hasinger et al. 2001) have discovered a population of absorbed point sources that outnumbers the Seyfert 1 galaxies by a factor of ∼4. These obscured AGN

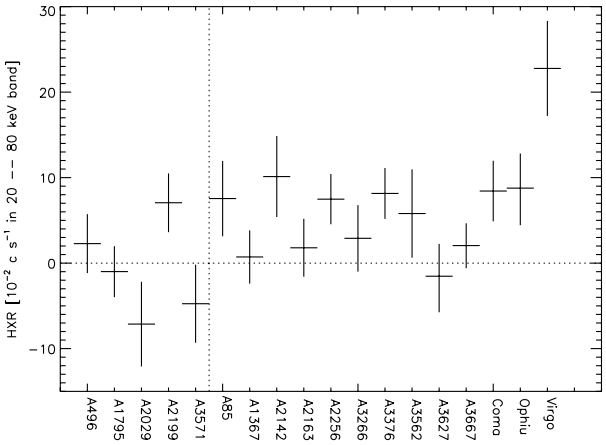


Fig. 3 NT signal and 1σ uncertainties in the PDS 20–80 keV band after subtraction of the contributions from the background, thermal emission, and AGN in the field, and accounting for uncertainties in these subtractions (Nevalainen et al. 2004). The *dotted vertical line* separates the relaxed clusters (*left*) from the rest (*right*)

are seen through an absorbing torus (with $N_{\rm H} = 10^{22}$–10^{25} atoms cm^{-2}; Risaliti et al. 1999) which hides them in the soft X-ray band. Thus, a robust estimate for their contribution was not available. However, the spectral and spatial distribution of the NT emission (see below) argues against significant contamination due to obscured AGN in the final Nevalainen et al. (2004) sample.

The thermal emission of clusters in the sample was modelled by Nevalainen et al. (2004) based on results of published analyses of the emission measured by the BeppoSAX MECS and XMM-Newton EPIC instruments. Since the PDS has no spatial resolution, the flux observed in the central region was extrapolated to the full PDS FOV using the appropriate single (or double) β models, after account was made for vignetting. They extrapolated the thermal emission model to the 20–80 keV band and compared the prediction with the PDS data, thus estimating the NT emission. The comparison showed that in ∼50% of the clusters the NT component was marginally detected at the 2σ level. Most of the significant detections were found in clusters which show signs of recent merger activity. Specifically, when clusters were divided into 'relaxed' or 'merger' groups, it was determined that the mean 20–80 keV NT emission in the latter group (4.8×10^{-2} counts s^{-1}) was ∼10 times higher than that of the 'relaxed' clusters (see Fig. 3). The level of systematic uncertainties due to background fluctuations is such that only the emission from clusters in the 'merger' group was found to be marginally significant (at $\sim 2\sigma$). Assuming a power-law shape for this excess with (photon) index of 2.0, it was deduced that the 20–80 keV luminosity per cluster is $4 \times 10^{43} h_{70}^{-2}$ erg s^{-1} (Nevalainen et al. 2004).

The residual emission of the individual clusters was too weak for meaningful spectral analysis; therefore, Nevalainen et al. (2004) co-added the data for the clusters, which amounted to a total of 560 ks. In order to obtain an estimate for the thermal contribution, the authors fitted data in the 12–20 keV band obtaining a best-fit temperature of ∼8 keV, consistent with the sample median. Extrapolating this model to the 20–80 keV band showed that there was indeed excess emission (see Fig. 4). A fit of the full 12–115 keV data to a thermal model resulted in a statistically unacceptable fit with an unrealistically high temperature of ∼26 keV.

A fit of the residual emission with a power-law model yielded a best-fit index $\Gamma = 2.8^{+0.3}_{-0.4}$ at 90% CL. Since a typical AGN index is ∼1.8, appreciable AGN emission was ruled out at the 98% CL. The high level of NT emission in the 12–20 keV band (∼50% of the total) is problematic, because in the central regions of the clusters in which NT emission was

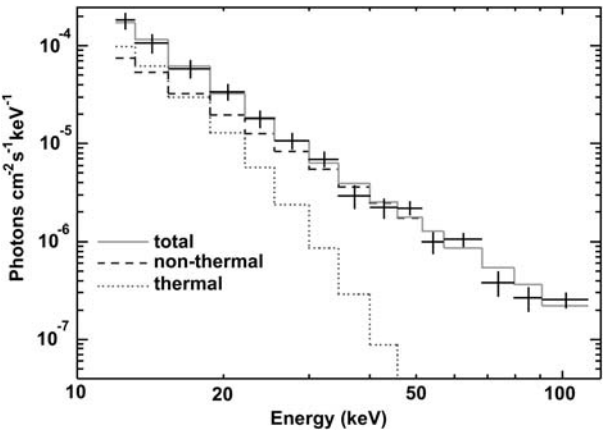


Fig. 4 The combined PDS spectrum of all the clusters not significantly affected by AGN (Nevalainen et al. 2004). *Lines* show the unfolded model components while *crosses* show the data and 1σ errors. The *dotted line* describes the thermal component, and the *solid line* shows the combined thermal and NT components. The *dashed line* indicates the best-fit power-law of $\Gamma = 2.8$

deduced, the MECS data typically allow the fractional contribution of such emission to be only a few percent. This seems to suggest that NT emission is extended with only a small fraction originating in the central region.

The indication from these statistical results for a connection between the cluster merger state and NT emission may be seen to support the scenario whereby electrons are accelerated by merger shocks. If so, the deduced power-law indices correspond to differential relativistic electron spectra with indices in the range $\mu = 3.8$–5.0. The implied radio synchrotron (energy) spectral indices $\sim$1.4–2, are consistent with the observed range.

3.5 Search for NT Emission with *ASCA*

ASCA had a fairly good sensitivity and low background up to about 10 keV, clearly too low for observing cluster NT emission. But there could also be detectable NT emission in groups of galaxies, in which galaxy interactions occur at relatively high encounter rates, and given that gas temperatures in groups are typically about 2 keV or less, an attempt was made to look for NT emission with *ASCA*. The first reported detection of excess emission, at energies above 5 keV, was made for the group HCG 62 (Fukazawa et al. 2001). The luminosity of the excess component was $\sim 4 \times 10^{41}$ erg s^{-1}, which was estimated to be some 20 times higher than the contribution of discrete X-ray sources. The hard excess was spatially extended to about $10'$ from the centre. Since the spectrum was too hard to be interpreted as thermal emission from the intra-group gas, a NT origin was thought to be more likely. A recent analysis of *Suzaku* XIS and HXD (see below) measurements of HCG 62 resulted only in an upper limit on NT emission (Tokoi et al. 2007), but at a level which does not exclude the ASCA result.

Nakazawa et al. (2007) carried out a search for NT emission from 18 groups of galaxies observed with *ASCA*, including HCG 62. They fitted the spectra below 2.5 keV with 2 temperature thermal (MEKAL) models and compared the data in the 4–8 keV band with the extrapolated thermal spectra. Excess fluxes are shown in Fig. 5 in terms of statistical significance. HCG 62 and RGH 80 show excess emission at $> 2\sigma$ CL, with excess emission thought to be likely also in NGC 1399. The residual spectra could be fitted either with a thermal model with $kT \geq 3$ keV or a power-law model with photon index fixed at 2. The 2–10 keV luminosity of the excess component is 10–30% of the thermal emission and 4–100 times stronger than the contribution from discrete X-ray sources. They concluded that both thermal and NT origins are acceptable from a statistical point of view.

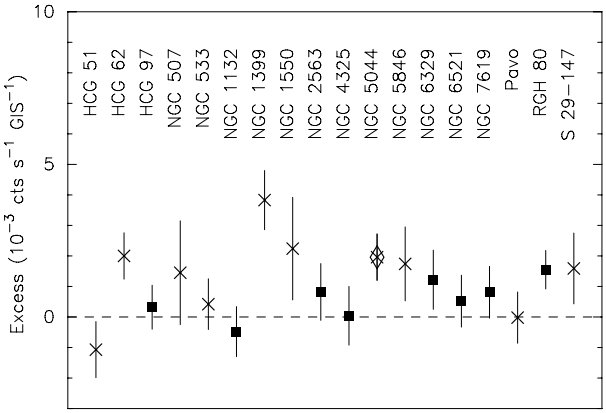


Fig. 5 *ASCA* measurement of excess emission in a sample of 18 groups of galaxies. The count rate in the 4–8 keV band is shown by the residual over the thermal model determined in the energy range below 2.5 keV. *Squares* denote groups for which the soft-band spectrum was fitted with single temperature model, while in the systems shown by stars two-temperature models were required. Error bars include statistical and systematic 1σ uncertainties. A thermal fit to the NGC 5044 spectrum, indicated with a *diamond*, was considered to be unacceptable. Adapted from Nakazawa et al. (2007)

3.6 Search for NT Emission with *Suzaku*

Suzaku, the 5th Japanese X-ray satellite launched in July 2005, carries hard X-ray detectors (HXD) along with the X-ray CCD instrument (XIS). The combined energy range is from 0.3 keV up to about 600 keV, and a wide-band coverage of X-ray and γ-ray sources is possible. The HXD system consists of 16 units of well-type phoswich detectors, surrounded by anti-coincidence shield scintillators (Takahashi et al. 2007; Kokubun et al. 2007). The well-type detector consists of a long active collimator made of BGO (a Bismuth-Germanium Oxyde), and the bottom of the well is equipped with Si PIN detectors and a GSO (a Gadolinium-Silicon Oxyde) scintillator. The PIN is 2 mm thick and sensitive from 8 keV to 50 keV; the GSO covers the 50–600 keV band. Typical effective areas are 160 cm^2 at 10 keV (PIN), and 330 cm^2 at 100 keV (GSO). The FOV is $34' \times 34'$ FWHM below 100 keV, limited by a phosphor bronze fine collimator installed in the well. At higher energies, the FOV becomes wider, up to $4.5^\circ \times 4.5^\circ$. The PIN detectors are cooled to -20°C by a thermal radiator connected with heat pipes.

The background level is very low for the PIN detector, since it has no accumulated effect of radio-activation in orbit. The measured background level in orbit is indeed lower than the level of the PDS (on BeppoSAX) by a factor of about 3, even though Suzaku's inclined orbit makes it more susceptible to the intense X-rays during passage through the radiation belt. This and the HXD narrow FOV make it the most sensitive instrument in the 10–60 keV band, among all previously flown hard X-ray detectors.

About 30 clusters have already been observed (as of April 2007) with Suzaku, including several for which the primary purpose is a search for NT emission. Meaningful constraints on NT emission were obtained for rather low temperature clusters. Most extensively observed is A 3376; these observations are described in some detail.

A 3376 This nearby ($z = 0.046$) merging cluster has an IC gas temperature of $\sim$4 keV. The relatively low temperature motivated the selection of the cluster as the prime target for

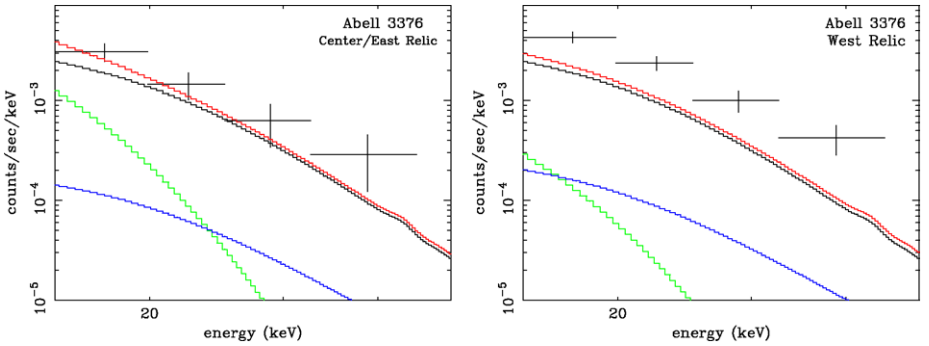


Fig. 6 Background subtracted spectra of A 3376 with the *Suzaku* HXD PIN detector (Kawano et al. 2007). *Left panel* is for the region including the centre and the east relic; the *right panel* is for the west relic. Upward along the left-hand ordinate, the *four solid curves* show estimated flux due to point-sources, cluster thermal emission, cosmic X-ray background contribution, and their sum, respectively

observations with the PIN detectors. The cluster is noted by two large radio relics, probably expanding over a Mpc scale in the east and west boundaries of the thermal emission. BeppoSAX observations resulted in a 2.7σ detection of a NT component. The Suzaku observations were carried out during two separate pointings in October and November 2005. The first ∼90 ks observation covered the central region which includes the east relic, and the second ∼ 103 ks was centred on the west relic region.

The background properties have been investigated by Kawano et al. (2007) and applied to the A 3376 data. Background data, i.e. data taken when the satellite was pointing to the dark Earth, were sorted by position of the satellite in Earth coordinates. This basically performs sorting with the cosmic-ray cut-off rigidity which is known to correlate well with the non X-ray background. Since passage through the radiation belt leaves enhanced emission afterward, the data in each Earth position were further divided into northward and southward satellite movements. The long-term trend of the background variation was also included in the estimation. Comparison with the predicted background rate and the observed dark Earth data indicates that this method gives 3.5% (1σ) error.

This method was applied to the observed data in the energy range 15–50 keV. Thermal cluster emission, estimated from the XIS measurements, CXB and point source contributions need to be subtracted. As shown in Fig. 6, the HXD PIN spectrum for the west relic suggests some level of excess emission. Including fluctuations of all these components, the flux of the west relic in the 15–50 keV band is $(6.3 \pm 1.8 \pm 6.2) \times 10^{-12}$ erg cm^{-2} s^{-1}, where the errors are statistical and systematic, respectively. Most of the systematic error, about 85%, is non-CXB; thus, further improvement in the background estimation is necessary in order to better determine the level of excess emission. The implied upper limit, 1.4×10^{-11} erg cm^{-2} s^{-1}, is consistent with the limit obtained from the *RXTE* measurements, and about 20% lower than the level reported from the *BeppoSAX* observation.

The west-relic region was jointly observed with the XIS instrument, and Kawano et al. also constrained the possible power-law flux based on the spectral fit. The obtained upper limit in the 4–8 keV region is extrapolated to 15–50 keV assuming a photon index of 1.8 with no scaling with the FOV, and it corresponds to about 20% of the hard X-ray upper limit. Of course, it is quite possible that the true NT flux is less than 3×10^{-12} erg cm^{-2} s^{-1} in the 15–50 keV band; however, if the emission is spatially extended then the XIS constraint still allows a hard flux of 2.4×10^{-11} erg cm^{-2} s^{-1} over the PIN field of view. Thus, an

extended hard X-ray emission (larger than 0.5°) remains a possibility. Using the 1.4 GHz measurement of the west relic, Kawano et al. deduced a lower limit $B_{\rm rx} > 0.1$ μG.

A 1060 and the Centaurus cluster These clusters are both relaxed nearby systems with $kT \sim 3.5$ keV; A 1060 is non-cD cluster, while the Centaurus cluster is dominated by a cD galaxy. The clusters were well studied by Kitaguchi et al. (2007). The observed data for the A 1060 and Centaurus cluster by the PIN instrument were accumulated for 28 and 26 ks, respectively. The latter authors applied the standard non-CXB subtraction and found that the blank-sky data yielded a fluctuation of 3–5% (1σ) in the energy range 10–50 keV. When the non-CXB flux was subtracted from the data, the data for both clusters showed agreement with the CXB above 20 keV; the A 1060 data even showed a deficit of the counts. In the energy range 10–20 keV, there remained an excess flux over the CXB level for both clusters. The flux was 4 or 8 times stronger than the CXB level for A 1060 and Centaurus, respectively. The cluster thermal emission was measured with *ASCA*, which had a wider FOV than the PIN, and the measured temperatures were 3.3 and 3.8 keV for the two clusters. These spectra with the intensity consistent with the *ASCA* measurement were compared with the PIN residual component. They found that the thermal emission can explain the PIN data quite well for both clusters, so only upper limits were derived on the NT emission.

3.7 *INTEGRAL* Observation of the Coma Cluster

A large fully coded field of view (8° × 8°), and good imaging capabilities with a PSF of 12′ FWHM, make it possible to construct hard X-ray cluster images below 300 keV with the IBIS/ISGRI coded mask instrument (Ubertini et al. 2003; Lebrun et al. 2003) aboard the *International Gamma-Ray Astrophysics Laboratory* (*INTEGRAL*, Winkler et al. 2003). Renaud et al. (2006) reported the analysis of a ∼500 ks observation of the Coma cluster with IBIS/ISGRI. The cluster was detected at a significance $\sim 10\sigma$ CL in the 18–30 keV energy band. The ISGRI image shows extended emission structure globally similar to that is seen by XMM-Newton below 10 keV (e.g. Arnaud et al. 2001; Neumann et al. 2003). In Fig. 7 we present two images of the Coma field in the 18–30 keV (left panel) and 30–50 keV (right panel) energy bands processed with the OSA-5 software, following a procedure similar to that described in Renaud et al. (2006), but for a longer (accumulated) ISGRI exposure of ∼ 940 ks ('good time'). The detected flux in the 18–30 keV band is consistent with the extrapolation of the flux measured by *XMM-Newton* below 10 keV assuming a model of a thin thermal plasma model with $kT = 8$ keV (see Renaud et al. 2006; Eckert et al. 2007); no statistically significant flux is seen in the higher energy band. From the observed images it appears that in order to detect or meaningfully constrain with *INTEGRAL* the presence of a NT X-ray component reported in *RXTE* and *BeppoSAX* observations of Coma, an integration time longer by a factor of ∼2–3 would be required.

3.8 Limits on Cluster γ-ray Emission

The relatively steep power-law shape of cluster NT emission (which is most precisely measured in the radio) directly implies that the detection of γ-ray emission from even nearby clusters will be very challenging. To estimate the predicted level of emission, consider the Coma cluster as an example; assuming no change in the relativistic electron power-law index, and the deduced 20–80 keV flux level (in Table 1), the cumulative flux at $\epsilon \geq 100$ MeV is $\sim 6 \times 10^{-9}$ cm^{-2} s^{-1}. Although substantially uncertain, flux at this level is close to the projected *GLAST* sensitivity threshold for an exposure time of 1 yr for the all sky survey.

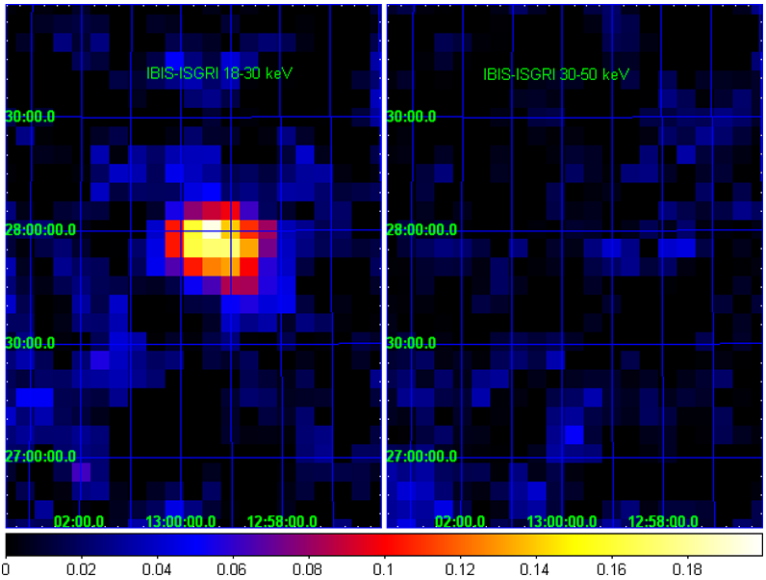


Fig. 7 Hard X-ray images of the Coma cluster from a 940 ks *IBIS-ISGRI* observation. The *left* and *right panels* show the *ISGRI* mosaics in the 18–30 keV, and 30–50 keV bands, respectively

The flux will be well below this level if the electron spectrum steepens even modestly at energies above the range deduced from radio measurements. Note that γ-ray emission could also have hadronic origin (from π^0 decay, following proton-proton interactions), but only very rough limits can be placed on this emission (e.g., Dermer and Rephaeli 1988).

A statistical upper limit on γ-ray emission from clusters was obtained by Reimer et al. (2003) from analysis of *EGRET* measurements. The emission in fields centred at 58 clusters observed between 1991 and 2000 was analysed. A co-added mean flux level was determined after accounting for the diffuse background contribution. The resulting upper 2σ limit on the mean cluster flux above (photon) energy of $\epsilon \geq 100$ MeV was found to be 6×10^{-9} cm^{-2} s^{-1}. Interestingly, this level is at the level predicted for Coma based on direct extrapolation of the deduced 20–80 keV flux. This *EGRET* upper limit was used by Bykov et al. (2000) and Petrosian (2001) to constrain models of hard X-ray emission components of leptonic origin; similarly Blasi and Colafrancesco (1999) used it to constrain the energetic proton energy density in IC space.

4 Discussion

It has been known all along that definite detection of NT X-ray emission in clusters is a challenging task. This is due to several factors, major among which are its intrinsically weak level—swamped as it is by the intense primary thermal emission—and the difficult to achieve requisite high sensitivity and low detector background in the 10–100 keV band. As has been emphasised, results of the search are not unequivocal, even for Coma and A 2256 for which detection (by *RXTE* and *BeppoSAX*) was claimed to be at moderately high level of statistical significance. This is mainly due to source confusion and a complete lack of spatial information. The search should continue (with *Suzaku* and future satellites), spurred

by its sound physical basis—the ubiquity of the CMB, and the observed radio synchrotron emission from relativistic electrons.

The expected level of NT emission of Compton origin depends steeply on the mean magnetic field strength in the central cluster region. Clearly, if this is typically a fraction of a μG, as has been deduced in the analyses described in the previous section, then prospects for more definite detection of NT emission—perhaps already with deep *Suzaku* measurements—are indeed good. However, significantly higher $B_{\rm fr}$ values—a few μG—are deduced from FR measurements; had these been meaningful estimates of the *volume-averaged* field strength, then definite detection of NT emission would have been seriously questioned.

As we have noted already (in Sect. 2.2), the apparent discrepancy between deduced values of $B_{\rm rx}$ and $B_{\rm fr}$ has been investigated at some length. These two field measures are quite different; the former is essentially a volume average of the relativistic electron density and (roughly) the square of the field, the latter is an average along the line of sight of the product of the field and gas density. All these quantities vary considerably across the cluster; in addition, the field is very likely tangled, with a wide range of coherence scales which can only be roughly estimated. These make the determination of the field by both methods considerably uncertain. Thus, the unsatisfactory observational status (stemming mainly from lack of spatial information) and the intrinsic difference between $B_{\rm rx}$ and $B_{\rm fr}$, do not allow a simple comparison of these quantities. Even if the large observational and systematic uncertainties, the different spatial dependences of the fields, relativistic electron density, and thermal electron density, already imply that $B_{\rm rx}$ and $B_{\rm fr}$ will in general be quite different. This was specifically shown by Goldshmidt and Rephaeli (1993) in the context of reasonable assumptions for the field morphology, and the known range of IC gas density profiles. They concluded that $B_{\rm rx}$ is indeed generally smaller than $B_{\rm fr}$. The implication is that prospects for detection of NT emission should not be based on the relatively high deduced values of $B_{\rm fr}$.

In conclusion, *RXTE* and *BeppoSAX* measurements have yielded appreciable evidence for power-law X-ray emission in four clusters. These results motivate further measurements and theoretical studies of NT phenomena on cluster and cosmological scales.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies".

References

M. Arnaud, N. Aghanim, R. Gastaud et al., Astron. Astrophys. **365**, L67 (2001)
P. Blasi, S. Colafrancesco, Astropar. Phys. **122**, 169 (1999)
A.M. Bykov, H. Bloemen, Yu.A. Uvarov, Astron. Astrophys. **362**, 886 (2000)
C.L. Carilli, G.B. Taylor, Annu. Rev. Astron. Astrophys. **40**, 319 (2002)
T.E. Clarke, J. Korean Astron. Soc. **37**, 337 (2004)
T.E. Clarke, P.P. Kronberg, H. Böhringer, Astrophys. J. **547**, L111 (2001)
L.P. David, A. Slyz, C. Jones et al., Astrophys. J. **412**, 479 (1993)
B. Dennison, Astrophys. J. **239**, L93 (1980)
C.D. Dermer, Y. Rephaeli, Astrophys. J. **329**, 687 (1988)
K. Dolag, S. Schindler, F. Govoni, L. Feretti, Astron. Astrophys. **378**, 777 (1999)
J.W. Dreher, C.L. Carilli, R.A. Perley, Astrophys. J. **316**, 611 (1987)
F. Durret, J. Kaastra, J. Nevalainen, T. Ohashi, N. Werner, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9313-8
D. Eckert, A. Neronov, T.J.-L. Courvoisier, N. Produit, Astron. Astrophys. **470**, 835 (2007)
T.A. Enßlin, C. Vogt, T.E. Clarke, G.B. Taylor, Astrophys. J. **597**, 870 (2003)
C. Ferrari, F. Govoni, S. Schindler, A.M. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9311-x

L. Feretti, D. Dallacasa, F. Govoni et al., Astron. Astrophys. **344**, 472 (1999)
Y. Fukazawa, K. Nakazawa, N. Isobe et al., Astrophys. J. **546**, L87 (2001)
R. Fusco-Femiano, D. Dal Fiume, L. Feretti et al., Astrophys. J. **513**, L21 (1999)
R. Fusco-Femiano, D. Dal Fiume, S. De Grandi et al., Astrophys. J. **534**, L7 (2000)
R. Fusco-Femiano, D. Dal Fiume, M. Orlandini et al., Astrophys. J. **552**, L97 (2001)
R. Fusco-Femiano, M. Orlandini, S. de Grandi et al., Astron. Astrophys. **398**, 441 (2003)
R. Fusco-Femiano, M. Orlandini, G. Brunetti et al., Astrophys. J. **602**, L73 (2004)
R. Fusco-Femiano, R. Landi, M. Orlandini, Astrophys. J. **624**, L69 (2005)
R. Fusco-Femiano, R. Landi, M. Orlandini, Astrophys. J. **654**, L9 (2007)
R. Gilli, G. Risaliti, M. Salvati, Astron. Astrophys. **347**, 424 (1999)
G. Giovannini, L. Feretti, New Astron. **5**, 335 (2000)
G. Giovannini, L. Feretti, T. Venturi, K.-T. Kim, P.P. Kronberg, Astrophys. J. **406**, 399 (1993)
G. Giovannini, M. Tordi, L. Feretti, New Astron. **4**, 141 (1999)
G. Giovannini, L. Feretti, F. Govoni, M. Murgia, R. Pizzo, Astron. Nachr. **327**, 563 (2006)
O. Goldshmidt, Y. Rephaeli, Astrophys. J. **411**, 518 (1993)
D.E. Gruber, Y. Rephaeli, Astrophys. J. **565**, 877 (2002)
F. Govoni, L. Feretti, Int. J. Mod. Phys. D **13**, 1549 (2004)
G. Hasinger, B. Altieri, M. Arnaud et al., Astron. Astrophys. **365**, L45 (2001)
M. Henriksen, Astrophys. J. **511**, 666 (1999)
M. Henriksen, R. Mushotzky, Astrophys. J. **553**, 84 (2001)
J.S. Kaastra, J.A.M. Bleeker, R. Mewe, Nucl. Phys. B **69**, 567 (1998)
J.S. Kaastra, R. Lieu, J. Mittaz et al., Astrophys. J. **519**, L119 (1999)
N. Kawano, K. Nakazawa, Y. Fukazawa et al., in *Proc. of 'The Extreme Universe in the Suzaku Era'* (2007 in press), conf. CD
K.-T. Kim, P.P. Kronberg, P.C. Tribble, Astrophys. J. **379**, 80 (1991)
T. Kitaguchi, K. Makishima, K. Nakazawa et al., in *Proc. of 'The Extreme Universe in the Suzaku Era'* (2007 in press), conf. CD
M. Kokubun, K. Makishima, T. Takahashi et al., Publ. Astron. Soc. Jpn. **59**, 53 (2007)
R. Landi, Ph.D. thesis, Bologna University, 2005
F. Lebrun, J.P. Leray, P. Lavocat et al., Astron. Astrophys. **411**, L141 (2003)
J.J. Mohr, B. Mathiesen, A.E. Evrard, Astrophys. J. **517**, 627 (1999)
S. Molendi, S. de Grandi, R. Fusco-Femiano et al., Astrophys. J. **525**, L73 (1999)
S. Molendi, S. De Grandi, M. Guainazzi, Astron. Astrophys. **392**, 13 (2002)
M. Murgia, F. Govoni, L. Feretti et al., Astron. Astrophys. **424**, 429 (2004)
K. Nakazawa, K. Makishima, Y. Fukazawa, Publ. Astron. Soc. Jpn. **59**, 167 (2007)
D.M. Neumann, D.H. Lumb, G.W. Pratt, U.G. Briel, Astron. Astrophys. **400**, 811 (2003)
J. Nevalainen, T. Oosterbroek, M. Bonamente, S. Colafrancesco, Astrophys. J. **608**, 166 (2004)
W.I. Newman, A.L. Newman, Y. Rephaeli, Astrophys. J. **575**, 755 (2002)
V. Petrosian, Astrophys. J. **557**, 560 (2001)
V. Petrosian, A.M. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9315-6
V. Petrosian, G. Madejski, K. Luli, Astrophys. J. **652**, 948 (2006)
V. Petrosian, A.M. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9327-2
O. Reimer, M. Pohl, P. Sreekumar, J.R. Mattox, Astrophys. J. **588**, 155 (2003)
M. Renaud, G. Belanger, J. Paul, F. Lebrun, R. Terrier, Astron. Astrophys. **453**, L5 (2006)
Y. Rephaeli, Astrophys. J. **212**, 608 (1977)
Y. Rephaeli, Astrophys. J. **227**, 364 (1979)
Y. Rephaeli, Mon. Not. R. Astron. Soc. **225**, 851 (1987)
Y. Rephaeli, in *Proc. of Heidelberg Int. Symp.* AIP Conf. Ser., vol. 558 (2001), p. 427
Y. Rephaeli, D.E. Gruber, Astrophys. J. **333**, 133 (1988)
Y. Rephaeli, D.E. Gruber, Astrophys. J. **579**, 587 (2002)
Y. Rephaeli, D.E. Gruber, Astrophys. J. **595**, 137 (2003)
Y. Rephaeli, D.E. Gruber, Astrophys. J. **606**, 825 (2004)
Y. Rephaeli, J. Silk, Astrophys. J. **442**, 91 (1995)
Y. Rephaeli, D. Gruber, R. Rothschild, Astrophys. J. **320**, 139 (1987)
Y. Rephaeli, D. Gruber, M. Ulmer, Astrophys. J. **429**, 554 (1994)
Y. Rephaeli, D. Gruber, P. Blanco, Astrophys. J. **511**, L21 (1999)
Y. Rephaeli, D.E. Gruber, Y. Arieli, Astrophys. J. **649**, 673 (2006)
M. Rossetti, S. Molendi, Astron. Astrophys. **414**, 41 (2004)
M. Rossetti, S. Molendi, astro-ph/0702417 (2007), electronic preprint only
G. Risaliti, R. Maiolino, M. Salvati, Astrophys. J. **522**, 157 (1999)
L. Rudnick, K. Blundell, Astrophys. J. **588**, 143 (2003)

J.S. Sanders, A.C. Fabian, Mon. Not. R. Astron. Soc. **381**, 1381 (2007)
C.L. Sarazin, Astrophys. J. **520**, 529 (1999)
C.L. Seeger, G. Westerhout, R.G. Conway, Astrophys. J. **126**, 585 (1957)
A. Simionescu, N. Werner, A. Finoguenov, H. Böhringer, M. Brüggen, Astron. Astrophys. (2007 submitted) (astro-ph/0709.4499)
T. Takahashi, K. Abe, M. Endo et al., Publ. Astron. Soc. Jpn. **59**, 35 (2007)
G.B. Taylor, R.A. Perley, Astrophys. J. **416**, 554 (1993)
M. Thierbach, U. Klein, R. Wielebinski, Astron. Astrophys. **397**, 53 (2003)
K. Tokoi, K. Sato, Y. Ishisaki et al., Publ. Astron. Soc. Jpn. **60** (2007 in press) (astro-ph/0711.1454)
P. Ubertini, F. Lebrun, G. Di Cocco et al., Astron. Astrophys. **411**, L131 (2003)
A. Valinia, M. Henriksen, M. Loewenstein et al., Astrophys. J. **515**, 42 (1999)
C. Vogt, T.A. Enßlin, Astron. Astrophys. **412**, 373 (2003)
M.A.G. Willson, Mon. Not. R. Astron. Soc. **151**, 1 (1970)
C. Winkler, T. Courvoisier, G. Di Cocco et al., Astron. Astrophys. **411**, L1 (2003)

Chapter 6
Observations of Extended Radio Emission in Clusters

C. Ferrari · F. Govoni · S. Schindler · A.M. Bykov · Y. Rephaeli

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9311-x

Abstract We review observations of extended regions of radio emission in clusters; these include diffuse emission in 'relics', and the large central regions commonly referred to as 'halos'. The spectral observations, as well as Faraday rotation measurements of background and cluster radio sources, provide the main evidence for large-scale intracluster magnetic fields and significant densities of relativistic electrons. Implications from these observations on acceleration mechanisms of these electrons are reviewed, including turbulent and shock acceleration, and also the origin of some of the electrons in collisions of relativistic protons by ambient protons in the (thermal) gas. Improved knowledge of non-thermal phenomena in clusters requires more extensive and detailed radio measurements; we briefly review prospects for future observations.

Keywords Galaxies: clusters: general · Galaxies: intergalactic medium · Radio continuum: general · Radiation mechanisms: non-thermal · Magnetic fields · Acceleration of particles

1 Introduction

In the last 10 years, the improved capabilities (sensitivity, spectral and spatial resolution) of multi-wavelength telescopes have allowed us to study in detail the formation and evolution

C. Ferrari (✉) · S. Schindler
Institute for Astro- and Particle Physics, University of Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria
e-mail: chiara.ferrari@uibk.ac.at

F. Govoni
INAF – Osservatorio Astronomico di Cagliari, Loc. Poggio dei Pini, Strada 54, 09012 Capoterra, Cagliari, Italy

A.M. Bykov
A.F. Ioffe Institute for Physics and Technology, 194021 St. Petersburg, Russia

Y. Rephaeli
School of Physics & Astronomy, Tel Aviv University, Tel Aviv 69978, Israel

of the largest gravitationally bound systems in the Universe, i.e. galaxy clusters. Following the hierarchical scenario of structure formation, massive clusters form through episodic mergers of smaller mass units (groups and poor clusters) and through the continuous accretion of field galaxies.

It has now been proven that major cluster mergers, with their huge release of gravitational binding energy ($\sim 10^{64}$ ergs), deeply affect the physical properties of the different components of clusters, i.e. the temperature, metallicity and density distribution of the thermal intracluster medium (ICM) emitting in X-rays (e.g., Buote 2002; Sauvageot et al. 2005; Ferrari et al. 2006a; Kapferer et al. 2006; Markevitch and Vikhlinin 2007), the global dynamics and spatial distribution of galaxies (e.g., Girardi and Biviano 2002; Ferrari et al. 2003, 2005; Maurogordato et al. 2008), as well as their star-formation rate (e.g., Gavazzi et al. 2003; Poggianti et al. 2004; Ferrari et al. 2006b). The typical signatures that allow to identify merging clusters from optical and X-ray observations are: a) substructures in the X-ray and optical surface densities (see Buote 2002, and references therein), b) non-Gaussian radial velocity distributions of cluster members (see Girardi and Biviano 2002, and references therein), c) patchy ICM temperature, pressure, entropy and metallicity maps (e.g., Finoguenov et al. 2005; Kapferer et al. 2006), d) sharp X-ray surface brightness discontinuities, accompanied by jumps in gas and temperature ("cold fronts", see, e.g., Markevitch et al. 2000), e) absent or disturbed cooling-cores (e.g., Markevitch et al. 1999), f) larger core radii compared to (nearly) relaxed clusters (Jones and Forman 1999). There are also indications that recent merging events lead to a depletion of the nearest cluster neighbours (e.g., Schuecker and Böhringer 1999). Additionally, deep radio observations have revealed the presence of diffuse and extended ($\sim$1 Mpc) radio sources in about 50 merging clusters. Their radio emission is not related to a particular cluster member, but rather to the presence of relativistic electrons (Lorentz factor $\gamma \gg 1000$) and weak magnetic fields (μG) in the intracluster space.

In this review, we focus on radio observations of this non-thermal component in galaxy clusters. We outline our current knowledge on the presence of non-thermal processes in the intracluster gas, and their physical connections with the thermodynamical evolution of large-scale structure. The relevance of the study of extended cluster radio emission for cosmology is pointed out. On smaller scales, there are only few indications of the possible presence of extended radio emission in galaxy groups. These systems host diffuse $\sim$1 keV gas called intragroup medium (IGM) (see, e.g., Mulchaey et al. 1996). Radio and hard X-ray emission possibly related to a non-thermal component of the IGM has been recently pointed out by Delain and Rudnick (2006) and Nakazawa et al. (2007) respectively. The existence of diffuse radio sources in galaxy groups has indeed to be tested with observations of higher sensitivity. Radio observations of the emission from individual radio galaxies are not treated here. For a discussion on cluster radio galaxies see the reviews by Feretti and Venturi (2002) and Feretti and Giovannini (2008). X-ray observations and simulations of the non-thermal component in clusters are reviewed by Rephaeli et al. (2008—Chap. 5, this issue) and Dolag et al. (2008—Chap. 15, this issue). The adopted cosmology is ΛCDM ($H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$).

2 Extended Radio Emission in Galaxy Clusters

The first detection of diffuse and extended radio emission in galaxy clusters dates back to 1959, when Large et al. mapped for the first time the Coma cluster at radio wavelengths, detecting an extended radio source (Coma C) at its centre. The existence of Coma C was later confirmed by Willson (1970), who compared the single dish data of Large et al. (1959) with

interferometric observations, and determined that the observed radio emission was diffuse and not associated with any cluster galaxy. From then on, high sensitivity radio observations have revealed in about 50 clusters the existence of diffuse non-thermal radio sources, not associated with active galaxies but with the ICM. The power-law radio spectrum[1] of this class of cluster sources indicates their synchrotron nature, and thus the presence of relativistic electrons (Lorentz factor $\gamma \gg 1000$) and magnetic fields ($\sim$0.1–1 μG) permeating the cluster volume.

While the thermal gas emitting in X-rays is present in all clusters, the detection of extended radio emission only in $\lesssim$10% of the systems indicates that the non-thermal plasma is not a common property of galaxy clusters. The very low surface brightness of diffuse cluster radio sources makes them difficult to detect with current radio telescopes. However, as we discuss here and in the following sections, our current knowledge suggests that the lack of radio emission in a high fraction of known clusters is not only related to a limited sensitivity of the current instruments,[2] but also to physical reasons, as non-thermal components over $\sim$1 Mpc scales are present only in the most massive merging clusters. Discriminating between these two effects is at present extremely difficult, and it will be one of the main goals of future radio observations (see Sect. 5).

The steep radio spectral index usually observed in diffuse cluster radio sources ($\alpha \gtrsim 1$) is indicative of ageing of the emitting particles. The steepening of the electron spectrum is a direct result of their Compton-synchrotron losses. The highest energy particles lose their energy more quickly. As a result, if cosmic rays are produced in a single event with a power law energy distribution

$$N(E)\,\mathrm{d}E = N_0 E^{-\delta}\,\mathrm{d}E \tag{1}$$

following the emergence of electrons from their sources (or acceleration sites), their spectrum steepens as result of the shorter energy loss time of high energy electrons. As a consequence the synchrotron spectrum falls off rapidly beyond a certain break frequency ν^*, which shifts gradually to lower frequencies. If instead particles are re-accelerated and/or continuously injected, as suggested to occur in cluster diffuse sources, the energy spectrum of relativistic particles may adopt more complex spectral shapes.

Apart from their common properties (nature of the emission, steep radio spectra), diffuse and extended radio sources in clusters differ in their physical properties, in particular: size, position in the host cluster, intensity of polarised signal, morphology and association to other cluster physical properties (e.g., dynamical state, presence of a cooling flow). A working definition, that is usually adopted, assigns cluster diffuse radio sources to three main classes: halos, relics and mini-halos. Very schematically:

- *radio halos* are extended ($\gtrsim$1 Mpc) diffuse radio sources at the centre of clusters, with a quite regular morphology, similar to the X-ray morphology of the system;
- *radio relics* have similar extensions and are also detected in merging clusters, but usually they are located in the cluster outskirts and they have an elongated morphology;
- *radio mini-halos* are smaller sources ($\lesssim$500 kpc) located at the centre of cooling flow clusters. They surround a powerful radio galaxy.

The main observational properties of these sources, useful to test their different formation scenarios (Sect. 3), are reviewed in more detail in the following (Sects. 2.1, 2.2 and 2.3).

[1] $S(\nu) \propto \nu^{-\alpha}$, see (3).

[2] For instance, the 3σ sensitivity limit for the NRAO VLA Sky Survey (NVSS) at 1.4 GHz, with a resolution of 45 arcsec, is $\sim$1.35 mJy/beam (Condon et al. 1998). The Very Large Array (VLA) is operated by the National Radio Astronomy Observatory.

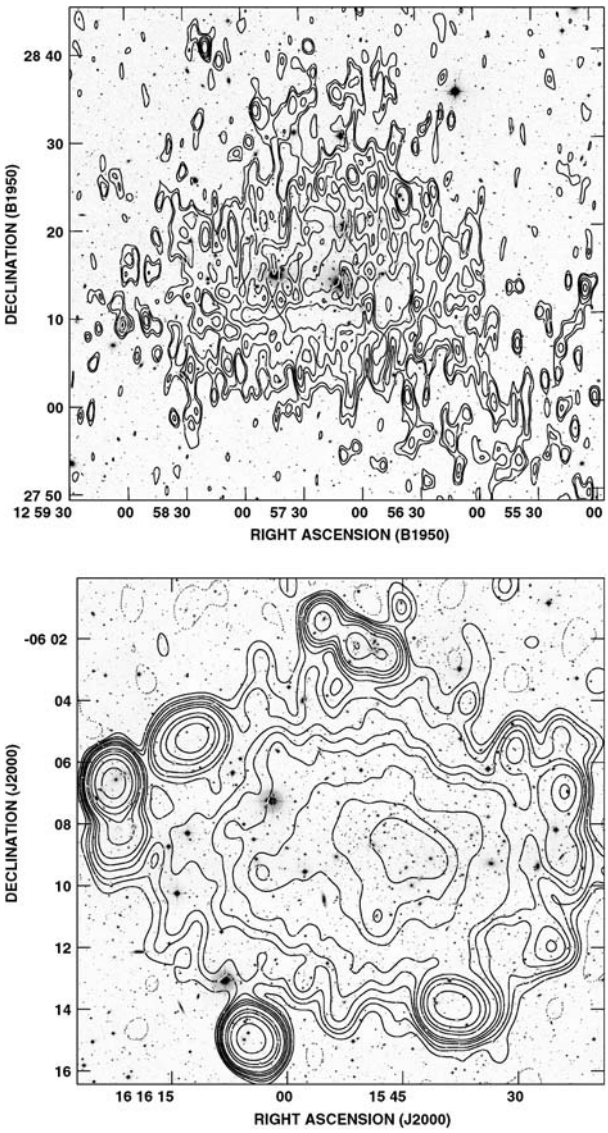


Fig. 1 *Top:* 90 cm contours of the radio halo in the Coma cluster ($z = 0.023$) are overlaid on the DSS optical image. Radio point sources have been subtracted (Feretti 2002). *Bottom:* 20 cm radio contours (Feretti et al. 2001) overlaid on the deep, R-band image (Maurogordato et al. 2008) of the galaxy cluster A 2163 ($z = 0.203$), hosting one of the most extended and powerful halos known so far

2.1 Radio Halos

Coma C is the prototype of the low surface brightness ($\sim\mu$Jy arcsec^{-2} at 1.4 GHz) and extended ($\gtrsim$1 Mpc) radio sources permeating the central volume of clusters, usually referred as "radio halos". Their radio morphology is quite regular (see Fig. 1) and their radio emission is unpolarised down to a few percent level. The first (and only) successful detection of polarised emission from a radio halo has been published recently by Govoni et al. (2005). One can find a compilation of most of the currently known radio halos in the recent review by Feretti and Giovannini (2008).

Spectral index studies of extended radio sources can give important hints on the energy spectrum of relativistic electrons and, due to the dependence of the synchrotron emissivity

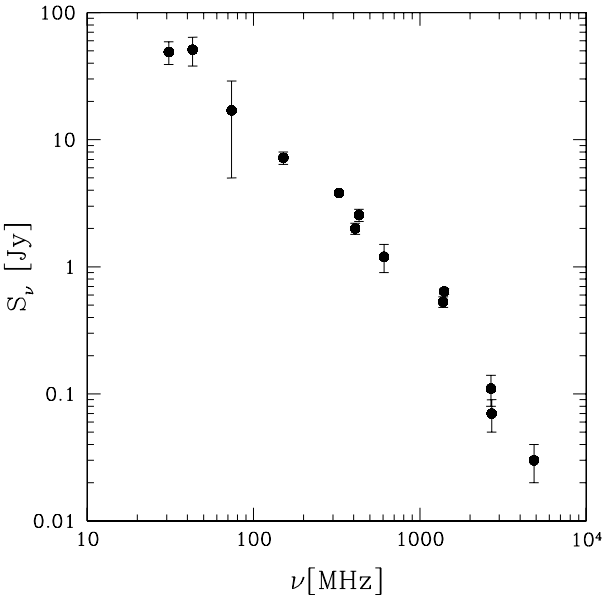


Fig. 2 Spectrum of the radio halo in the Coma cluster (Coma C). A steepening in the spectrum stands out clearly at $\nu > 1$ GHz (adapted from Thierbach et al. 2003)

on the magnetic field intensity (3), on the magnetic field distribution (Brunetti et al. 2001). In recent years, many observational efforts have thus been devoted to multi-frequency observations of radio halos, in order to get more and more accurate determinations of the integrated radio spectrum and, possibly, of spatially resolved spectral index maps. These studies are limited however by the capability of current instruments to do multi-frequency observations at the sensitivity required for studying radio halos ($\sim$mJy–μJy arcsec^{-2} going from the MHz to the GHz range). In a few cases, a steepening of the halo spectrum at high frequency (as in Fig. 2 in the case of the Coma cluster) has been detected (A 2319: Feretti et al. 1997; Coma: Thierbach et al. 2003; A 754: Bacchi et al. 2003; A 3562: Giacintucci et al. 2005). Indications that the spectral index steepens radially with the distance from the cluster centre have been pointed out by Giovannini et al. (1993) in Coma, and Feretti et al. (2004a) in A 665 and A 2163. More recently, Orrù et al. (2007) have shown that the radio halos in A 2744 and A 2219 have a mean spectral index, averaged over the whole cluster, without a clear radial steepening, but with a very patchy structure. Very interestingly, their radio/X-ray comparison shows flatter spectral indexes in regions characterised by higher ICM temperature.

Current observational results suggest other strong connections between the physical properties of radio halos and of their host clusters. All radio halos discovered up to now are at the centre of clusters with signatures of a disturbed dynamical state and without a cooling core. However, not all merging clusters host a radio halo. The detection rate of radio halos is actually quite low: 5% in a complete cluster sample at the detection limit of the NVSS, which grows to $\sim$35% when only the most luminous X-ray clusters are considered ($L_{X[0.1-2.4\ \mathrm{keV}]} > 0.6 \times 10^{45}\ \mathrm{h}_{70}^{-2}\ \mathrm{erg\,s^{-1}}$) (Giovannini et al. 1999; Giovannini and Feretti 2002). The fact that radio and X-ray properties of clusters are connected is also suggested by a close similarity of the morphology of radio halos and the X-ray emission of their host clusters. This has firstly been revealed in a qualitative way (Deiss et al. 1997; Feretti 1999; Liang et al. 2000), and afterwards quantitatively confirmed by the relation between

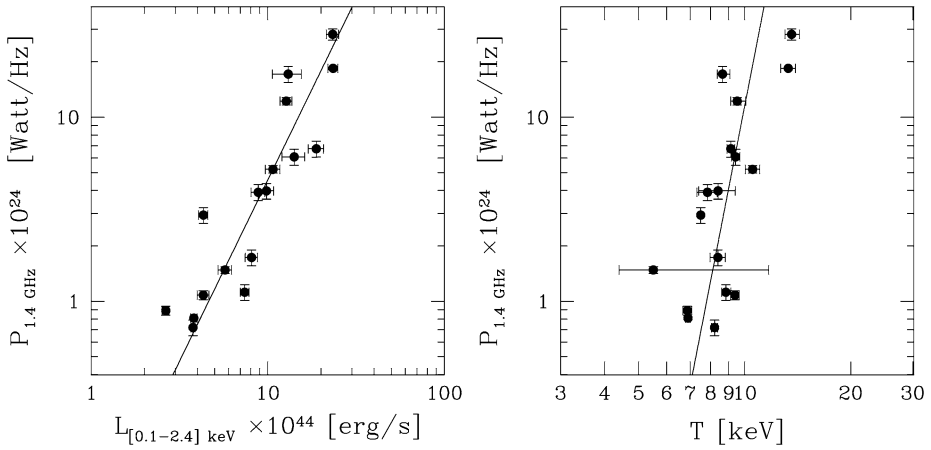


Fig. 3 Radio power at 1.4 GHz of giant ($\gtrsim$1 Mpc) radio halos vs. a) *left:* cluster X-ray luminosity, and b) *right:* cluster X-ray temperature (adapted from Cassano et al. 2006)

the point-to-point surface brightness of the cluster radio and X-ray emission (Govoni et al. 2001c; Feretti et al. 2001).

Additionally, a strong correlation has been pointed out between the radio power (P_ν) of halos and the X-ray luminosity (L_X) of their host clusters (e.g., Liang et al. 2000; Giovannini and Feretti 2002; Enßlin and Röttgering 2002; Cassano et al. 2006; see left panel of Fig. 3). A relation with a much larger scatter between radio power and X-ray temperature of the ICM (T_X) has also been suggested (e.g., Colafrancesco 1999; Liang et al. 2000) (see right panel of Fig. 3). Since both the X-ray luminosity and temperature of clusters correlate with mass (e.g., Neumann and Arnaud 1999; Neumann and Arnaud 2001), the observed P_ν–L_X and P_ν–T_X relation could reflect a dependence of the radio halo luminosity on the cluster mass, with interesting implications on the theoretical models of cosmic ray production (see Sect. 3). Current results suggest $P_{1.4\ \mathrm{GHz}} \propto M^a$, with $a = 2.3$ or larger, depending on the methods applied to estimate the cluster mass (see Feretti and Giovannini 2008 and references therein).

Recently, Cassano et al. (2007) pointed out that the fraction of the radio emitting cluster volume significantly increases with the cluster mass. This break of self-similarity can give important constraints on the physical parameters entering the hierarchical formation scenario, since it suggests that the distributions of the magnetic field and relativistic electrons change with the cluster mass.

2.2 Radio Mini-Halos

The so-called "radio mini-halos" (Fig. 4) differ from the above described radio halos not only because of their smaller size (few 10^2 kpc), as their name suggests, but also in the typical properties of their host clusters. Actually, mini-halos are diffuse radio sources with a steep spectral index, which are found around powerful radio galaxies at the centre of cooling core clusters. The total size of mini-halos is comparable to that of the cooling region. Since major mergers are able to disrupt (or at least disturb) cluster cooling flows (e.g., Buote and Tsai 1996; Gómez et al. 2002), the main physical difference between giant and mini-halos is that they are hosted in clusters with and with no evidence of *major* mergers respectively.

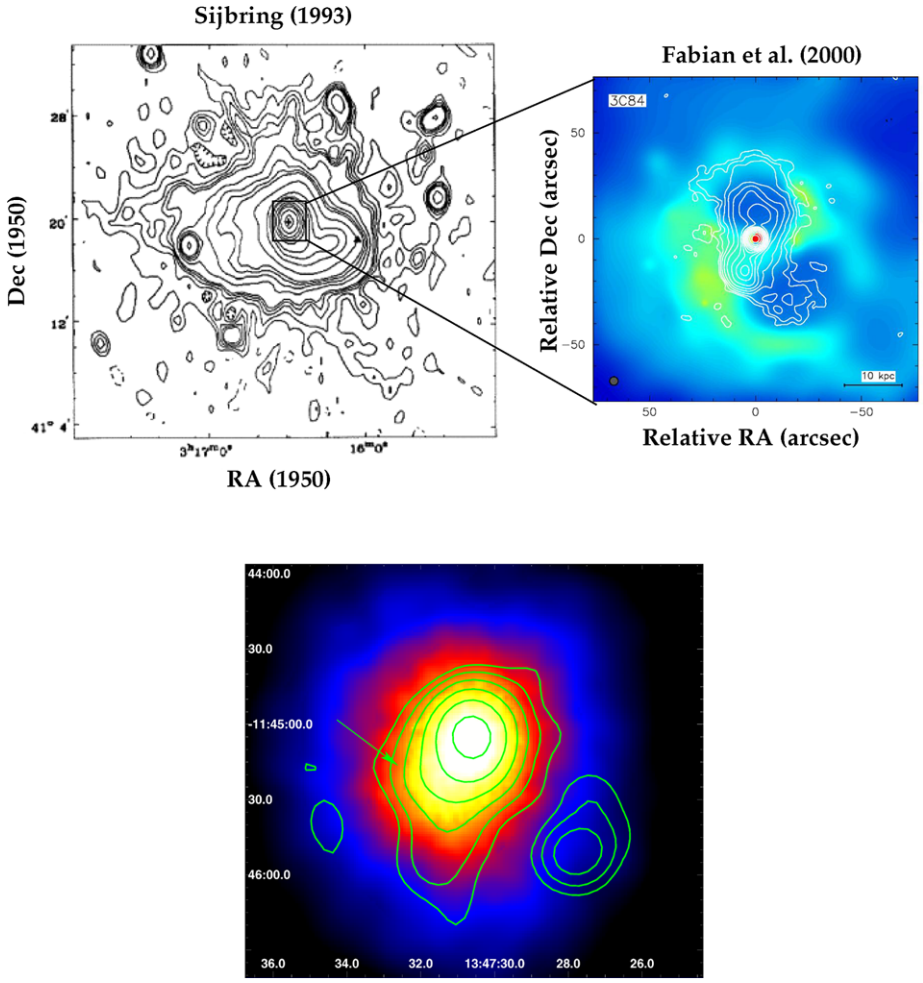


Fig. 4 *Top:* 327 MHz map of the mini-halo in the Perseus Cluster ($z = 0.018$). The source is centred on the position of the cD galaxy NGC 1275 (indicated with a *cross*). The *inset* shows radio contours overlaid on the X-ray image of the central $\sim 1'$ region of Perseus. The holes evident in the X-ray emission are due to subsonic expansion of the buoyant radio lobes of the central radio galaxy 3C 84 (adapted from Sijbring 1993 and Fabian et al. 2000). *Bottom:* 1.4 GHz contours of the radio mini-halo (the most distant ever detected) in the galaxy cluster RXJ 1347.5–1145 ($z = 0.451$), superimposed on the *XMM-Newton* image of the cluster. The green arrow indicates an elongation in the radio mini-halo morphology, corresponding to a sub-structure detected in X-rays (adapted from Gitti et al. 2007a)

However, recent results revealed the existence of two cooling flow clusters with signatures of merging activity in the central region and hosting a radio mini-halo: A 2142 (Giovannini and Feretti 2000) and RXJ 1347.5–1145 (Gitti et al. 2007a, see bottom panel of Fig. 4). Contrary to what expected in relaxed systems, both the clusters are dominated by two brightest cluster galaxies (BCGs). In A 2142, the central cooling flow has been disturbed but not destroyed by an unequal merger, observed 1–2 Gyr after the initial core crossing (Markevitch et al. 2000). The cooling flow in RXJ 1347.5–1145 is one of the most massive ever detected, suggesting a relatively long interval of time in which the central part

of the cluster has evolved undisturbed to a nearly relaxed state (Gitti et al. 2007b). The X-ray analysis of Gitti et al. (2007b) reveals however a sub-structure in the south-east part of the cluster, corresponding to an elongation in the radio mini-halo morphology (see bottom panel of Fig. 4). Indications of possible *minor* mergers have been detected also in other clusters hosting radio mini halos (Perseus: Ettori et al. 1998; Furusho et al. 2001; A 2390: Allen et al. 2001; A 2626: Mohr et al. 1996). In these cases, however, the merging substructures are located well outside the diffuse radio source.

Several radio halos have been discovered in radio surveys (e.g., NVSS, Giovannini et al. 1999), where their relatively large beam provides the necessary signal-to-noise ratio to spot these elusive sources. Due to their extremely low surface brightness and large angular extent radio halos are actually best studied at low spatial resolution. Unfortunately, due to the combination of their smaller angular size and the strong radio emission of the central radio galaxy, the detection of a mini-halo requires a much higher dynamic range and resolution than those in available surveys, and this complicates their detection. As a consequence, our current observational knowledge on mini-halos is limited to less than ten known sources (PKS 0745–191: Baum and O'Dea 1991; Perseus: Burns et al. 1992; A 2142: Giovannini and Feretti 2000; A2390: Bacchi et al. 2003; A 2626: Rizza et al. 2000; Gitti et al. 2004; RXJ 1347.5–1145: Gitti et al. 2007a). This, together with the peculiar properties of A 2142 and RXJ 1347.5–1145, and the possible connection between radio mini-halos and minor cluster mergers, requires better statistics to test the current theoretical models on the origin of their radio emission, which are discussed in Sect. 3.

2.3 Radio Relics

As clearly stated by Kempner et al. (2004) there is quite a lot of confusion in the literature when speaking about "radio relics". This definition is actually adopted for at least three different kinds of radio sources in galaxy clusters, characterised by significantly different observational properties. Certain features are in common for all the different kinds of relic sources, such as their steep radio spectrum ($\alpha \gtrsim 1$) and a general filamentary morphology.

A first group of sources (see an example in the top panel in Fig. 5) has typical sizes of several 10 kpc and low/intermediate polarisation intensity ($\lesssim$20%). They are generally located in the central cluster regions, close to the cD galaxy, often showing an anti-correlation with the ICM density. Actually, in some clusters, relic as well as AGN radio emission has been detected inside holes in the central X-ray emission of the thermal gas (see for instance the inset in the top panel of Fig. 4). These cavities are actually related to the cyclic outburst activity of the central AGN. The cases in which no radio emission has been detected, or has been revealed only by low frequency radio observations (e.g., the Perseus outer cavities discovered by Fabian et al. 2002), are due to buoyant old radio lobes, whose spectrum is too steep to be detected in the GHz range. A more detailed discussion on these "radio ghosts" and X-ray cavities in clusters can be found in a recent review by McNamara et al. (2008).

Both the second and the third class of "cluster relics" (middle and bottom panel of Fig. 5) are strongly correlated with the ICM properties. They are commonly found in merging clusters, and, in some cases, a spatial correlation with shocks in the thermal gas has been pointed out (e.g., Kassim et al. 2001). Both of these classes of sources do not have a likely parent radio galaxy nearby, are generally polarised at the level of about 10–30% and located in the cluster periphery. While the (rare) objects in the second group are characterised by intermediate linear scales compared to the other two (with typical sizes of $\sim 10^2$ kpc, e.g., Slee et al. 2001), the most extended radio emission among the "relic sources" comes from the third class.

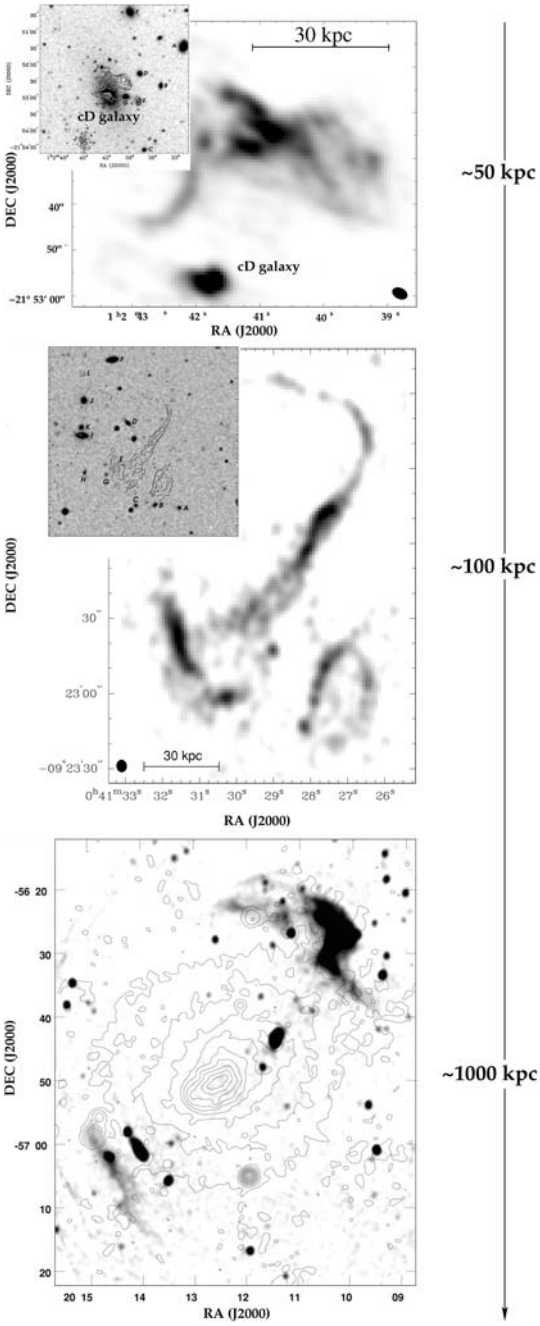


Fig. 5 Examples of different diffuse radio sources in clusters classified as "relics" in the literature. From top to bottom, an "AGN relic", a "Phoenix" and a "Radio Gischt" (see Kempner et al. 2004 and Sects. 2.3 and 3). *Top panel:* VLA image at 1.4 GHz of the "AGN relic" source in A 133 ($z = 0.056$), close to the radio emitting cD galaxy at the cluster centroid. In the inset, radio contours are superimposed on the optical, DSS-2 image of the area (adapted from Slee et al. 2001). *Middle panel:* VLA image at 1.4 GHz of the "Phoenix" source in the periphery of A 85 ($z = 0.055$) (see also Fig. 11 in Slee et al. 2001). As before, the inset shows radio contours superimposed on the optical DSS-2 image of the region around the radio source (adapted from Slee et al. 2001). *Bottom panel:* X-ray contours (ROSAT data, 0.1–2.4 keV energy band) overlaid on the 843 MHz Molonglo Observatory Synthesis Telescope (MOST) image of A 3667 ($z = 0.056$) (from Röttgering et al. 1997)

These giant relics, with extensions ranging from a few $\sim 10^2$ kpc to 10^3 kpc, have been detected in merging clusters both with and without cooling cores. The major axis of their elongated structure is generally nearly perpendicular to the direction of the cluster radius. Some of the most extended and powerful giant relics are located in clusters with central radio

halos (e.g., A 2256: Clarke and Enßlin 2006). In a few cases, two symmetric relics have been observed, as in the bottom panel of Fig. 5 (A 3667: Röttgering et al. 1997; Johnston-Hollitt et al. 2002; A 3376: Bagchi et al. 2006; A 1240: Kempner and Sarazin 2001). More "exotic" giant radio relics have also been discovered, such as sources located in poor clusters (Abell S 753: Subrahmanyan et al. 2003), far away from the cluster centre (A 786: Giovannini and Feretti 2000), or even in intracluster filaments of galaxies (ZwCl 2341.1+0000: Bagchi et al. 2002). Relics elongated from the cluster centre to the periphery (A 115: Govoni et al. 2001b), or with a circular shape (A 1664: Feretti 2005) have also been detected. Similarly to radio halos (Sect. 2.1), the detection rate of giant radio relics at the sensitivity limit of NVSS is ∼6% (Giovannini and Feretti 2002), and the relic radio power correlates with the cluster X-ray luminosity (Feretti 2002), even though with a larger scatter.

3 Origin of Radio Emitting Particles

Based on the observational results summarised in Table 1 and in the previous sections, and on the theoretical models reviewed, for instance, by Brunetti (2004), Blasi et al. (2007) and Dolag et al. (2008), we have now a formation scenario for the different diffuse and extended radio sources in clusters. The current theories on the origin of the non-thermal component in galaxy clusters will be the starting point for new observational studies with the next generation radio telescopes (Sect. 5).

Giant radio halos and relics are the most spectacular radio sources in clusters, and, as stated above, their synchrotron spectrum indicates the presence of cosmic rays that gyrate around magnetic field lines, frozen in the ICM. Therefore, relativistic particles cannot stream out from the gravitational field of the cluster, but they can still diffuse along magnetic field lines. It has been shown however (Völk et al. 1996; Berezinsky et al. 1997; Völk and Atoyan 1999) that typical relativistic electrons in radio halos and relics (with $\gamma \sim 1000$–5000) have diffusion times which are longer than the Hubble time. They could therefore be simply diffused over cluster scales from one or more active radio galaxies (Jaffe 1977; Rephaeli 1977). However, the steep radio spectra of these sources indicate short lifetimes for the radiating particles ($\sim 10^8$ yr), which lose energy not only via synchrotron emission, but also due to interactions with the Cosmic Microwave Background (CMB) photons (via Compton scattering emission) and with the ICM (via Coulomb interactions and Bremsstrahlung emission). The main radiative losses of electrons are due to Compton scattering of the CMB, and synchrotron emission; the former process dominates for $B < 3$ μG (the field equivalent of the CMB energy density). The radiative lifetime of a relativistic lepton with a Lorentz factor

Table 1 Main observational properties of the different sources of diffuse radio emission in galaxy clusters. Note that linear polarisation levels of ∼20–40% have been detected in filamentary structures of the radio halo in A 2255 by Govoni et al. (2005)

Type	Position	Size	α	Polarisation	Example
Halo	Centrally peaked	$\gtrsim$Mpc	$\gtrsim$1	<few %	Coma
Giant relic	Peripheral	∼Mpc	$\gtrsim$1	∼10–30%	Abell 3667
Mini-halo	Centrally peaked	$\lesssim$0.5 Mpc	$\gtrsim$1.5	<few %	Perseus
Phoenix	Peripheral	$\sim 10^2$ kpc	$\gtrsim$1.5	∼10–30%	Abell 85
AGN relic	Close to the host galaxy	few ×10 kpc	$\gtrsim$1.5	$\lesssim$20%	Abell 133

$\gamma < 10^8$ is thus approximately given by (e.g., Longair 1981; Meisenheimer et al. 1989)

$$\tau \approx 2 \times 10^{12} \gamma^{-1} \left[(1+z)^4 + \left(\frac{B}{3.3 \mu \mathrm{G}} \right)^2 \right]^{-1} \text{ years.} \quad (2)$$

Since the expected diffusion velocity of the relativistic electrons is of the order of 100 km s^{-1} (Alfvén speed), cosmic rays do not have the time to propagate over the Mpc-scales of giant cluster radio sources. This excludes the hypothesis that relativistic electrons are produced at a localised point, and requires *in situ* acceleration mechanisms. Basically, two classes of models have been proposed:

– the *primary models*, which predict that electrons are accelerated by shocks and/or turbulence induced during cluster mergers. Shocks can (re-)accelerate particles via Fermi-I processes (Enßlin et al. 1998) or adiabatic compression of fossil radio plasma (Enßlin and Gopal-Krishna 2001); turbulence via stochastic, Fermi-II processes (Brunetti et al. 2001) or magnetohydrodynamic (MHD) waves (Alfvén waves: Brunetti et al. 2004; magnetosonic waves: Cassano and Brunetti 2005);
– the *secondary models*, in which relativistic electrons are continuously injected by hadronic collisions between the thermal ions of the ICM and relativistic protons, the latter (characterised by significantly longer lifetime compared to relativistic electrons) having been accelerated during the whole cluster history (e.g., Dennison 1980; Blasi and Colafrancesco 1999; Dolag and Enßlin 2000).

The observational properties of radio halos and relics (see Sects. 2.1 and 2.3) are more in favour of primary models. The strongest point leading to this conclusion is the fact that diffuse and extended radio emission has been detected up to now only in merging clusters. A strong connection between cluster mergers and cosmic ray production is required in primary models, and is not expected in secondary models. In this respect, the fact that halos and relics are quite rare in clusters is again disfavouring the hadronic collision hypothesis, based on which we should expect electron acceleration to be possible in all galaxy clusters.

Since the shape of the synchrotron spectrum depends on the last acceleration phase of cosmic rays, detailed studies of the spectral index distribution in radio halos and relics provide important information on acceleration mechanisms acting in clusters. Primary models predict a radial steepening and a complex spatial distribution of the spectral index α, due to the existence of a) a maximum energy to which electrons can be accelerated ($\gamma < 10^5$, Blasi 2004, and references therein) and b) different re-acceleration processes in different cluster regions. Secondary models assume that cosmic ray protons are accelerated during structure formation over cosmological epochs and accumulated in clusters. The collision of these protons with the thermal ICM would continuously inject electrons, resulting in a spectral index distribution unrelated to the intracluster magnetic field strength, and thus not dependent on the position in the cluster. The radial spectral steepening and/or the patchy structure of spectral index maps observed in several radio halos (Sect. 2.1) are clearly favouring primary models.

The hadronic collision hypothesis predicts power-law spectra flatter than primary models ($\alpha \lesssim 1.5$) and magnetic field values higher than a few μG. Observational results are controversial concerning these points, due to the observed intermediate values of the radio spectral index ($\alpha \sim 1$–1.5), and the widely differing estimates for mean intracluster magnetic field values (Sect. 4). Finally, emission of gamma-rays is expected in secondary models (Blasi et al. 2007), a challenging point to be tested observationally (Rephaeli et al. 2008—Chap. 5, this issue). On the other hand, it has been suggested (Bykov et al. 2000; Miniati 2003) that

the detection of gamma-ray emission from clusters may not necessarily reflect the hadronic origin of cosmic rays, since it could be related to the Compton scattering of CMB photon from shock-accelerated, intracluster electrons.

Given our current observational and theoretical knowledge, cosmic rays in giant radio relics (bottom panel of Fig. 5) are most likely originating from Fermi-I diffuse shock acceleration of ICM electrons (e.g., Hoeft and Brüggen 2007; Bykov et al. 2008—Chap. 7, this issue). These radio sources would therefore trace the rim of shock fronts resulting from cluster mergers in the ICM, and they have been named "radio gischt"[3] by Kempner et al. (2004). Firstly, this hypothesis is in agreement with the morphology and the position of most of the detected giant relics, which appear as elongated, sometimes symmetric, radio sources in the cluster periphery, where we expect to find arc-like shock fronts resulting from major cluster mergers (e.g., Schindler 2002). Secondly, the quite strong linear polarisation detected in giant relics would be in agreement with the model prediction of magnetic fields aligned with the shock front. Based on some observational results, however, a clear association between shocks and giant radio relics is not always straightforward. This is true in the case of the "exotic" giant relics mentioned in Sect. 2.3 (e.g., those with circular shapes, or located in intracluster filaments). Additionally, Feretti and Neumann (2006) did not detect a shock wave corresponding to the radio relic in the Coma cluster. They suggested that, similarly to radio halos (see below), the radio emission of this relic source is instead related to turbulence in the ICM. Currently, the main observational limitation to test the origin of giant radio relics comes from X-ray data. The sensitivity of X-ray instruments is not high enough to detect shock waves in the external regions of clusters, where the gas density and thus the X-ray surface brightness are very low, and where most of radio relics have been detected (see for instance the radio relic found in A 521 by Ferrari (2003); see also Ferrari et al. 2006a; Giacintucci et al. 2006).

The second class of relic sources pointed out in Sect. 2.3 (middle panel in Fig. 5), characterised by smaller sizes than giant relics ($\sim 10^2$ kpc vs. 10^3 kpc), are most likely originating from adiabatic compression in cluster shocks of fossil radio plasma, released by an AGN whose central engine has ceased to inject fresh plasma $\lesssim$1 Gyr ago (Enßlin and Gopal-Krishna 2001). The old non-thermal electrons, that would be undetectable at high ($\sim$GHz) frequencies, are actually re-energised by the shock. This class of sources are therefore also called "Phoenix" (Kempner et al. 2004). The main physical difference with giant relics is related to the fact that, in the latter, shocks accelerate thermal ICM electrons to relativistic velocities through Fermi-I processes, while, in the case of Phoenix sources, the shock waves energise the relativistic plasma by adiabatic compression. The sound speed inside the fossil radio plasma is actually so high that shock waves cannot penetrate into the radio cocoons. These sources are rare because they require shocks and fossil plasma in the same region of the cluster. Moreover, adiabatic compression is efficient in re-accelerating electrons only if the time elapsed since the last AGN activity is not too long, i.e. less than about 0.1 Gyr in the cluster centre and less than about 1 Gyr in the periphery. All this also explains why we detect Phoenix sources in the external regions of clusters.

Among the sources called "radio relics" in the literature, only the smallest (several tens of kpc) are real "AGN relics" (top panel in Fig. 5). They are extinct or dying AGNs, in which the central nucleus has switched off, leaving the radio plasma to evolve passively (e.g., Murgia et al. 2005; Parma et al. 2007). Their spectrum becomes steeper and steeper, making the source more and more difficult to be detected at high frequencies, until it disappears completely (see Sect. 2.3). Due to the short radiative lifetime of their electrons ($\sim 10^7$–10^8 yrs),

[3]German for the spray on the tops of ocean waves.

these sources are usually located close to their host galaxy, which did not have time enough to move far away in the cluster potential.

In the case of giant radio halos, spectral index maps show no evidence of flattening at the location of shocks detected in X-rays (A 665: Feretti et al. 2004a; Markevitch and Vikhlinin 2001; A 2744: Orrù et al. 2007; Kempner and David 2004). This agrees with theoretical results showing that shocks in major mergers are too weak to produce relativistic particles uniformly over the whole central $\sim$1 Mpc area of clusters (Gabici and Blasi 2003). Although it cannot be excluded that shock acceleration may be efficient in some particular regions of a halo (e.g., Markevitch et al. 2005), it has been suggested that cluster turbulence generated by cluster mergers may efficiently accelerate electrons in the cluster volume (e.g., Cassano and Brunetti 2005). The observed steepening of the spectral index with the distance from the cluster centre, and the few available spectral index maps showing flatter spectra in the regions influenced by merger processes (Sect. 2.1) support the scenario that ICM turbulence supplies the energy for the radiating electrons.

However, if the predictions of primary models better agree with the observational results, the merging event cannot be solely responsible for electron re-acceleration in giant radio halos and relics, because $\gtrsim$40% of clusters show evidence of a disturbed dynamical state (Jones and Forman 1999), while only $\lesssim$10% possess radio halos and/or relics. As we have seen in Sect. 2.1, the power of observed radio halos P_ν seems to correlate with the mass M of their host cluster. The energy available to accelerate relativistic particles during cluster mergers is a fraction of the gravitational potential energy released during the merging event, that in turn scales as $\sim M^2$. The P_ν–M relation could thus suggest that only the most massive mergers are energetic enough to efficiently accelerate cosmic rays (Buote 2001). A recent model by Cassano and Brunetti (2005) is in agreement with this conclusion, showing that only massive clusters can host giant radio halos. The probability to form these extended radio sources increases drastically for cluster masses $M \geq 2 \times 10^{15}\ M_\odot$ since the energy density of the turbulence is an increasing function of the mass of the cluster. Based on the scenario of hierarchical structure formation, massive clusters result from a complex merging history, during which each cluster–cluster collision could have contributed to provide energy for cosmic ray acceleration.

Finally, as we have seen in Sect. 2.2, radio mini-halos have also been observed in clusters. They are located at the centre of cooling flow clusters and surround a powerful radio galaxy. Similarly to giant radio halos and relics, the electrons in radio mini-halos have short radiative lifetimes due to the high magnetic fields present in cooling cores (Taylor et al. 2002). The observed radio emission is thus not due to the radio lobes of the central AGN. Unlike the giant sources, mini-halos are typically found in clusters not disturbed by major mergers (Sect. 2.2). Again, two possible classes of models have been proposed. Relativistic electrons could have again an hadronic origin (Pfrommer and Enßlin 2004). Or they could be a relic population of (primary) relativistic electrons re-accelerated by MHD turbulence, with the necessary energy supplied by the cooling flow (Gitti et al. 2002). The re-acceleration model by Gitti et al. (2002) has been successfully applied to two cooling flow clusters (Gitti et al. 2002, 2004). The observed correlation between the mini-halo and cooling flow power has also given support to a primary origin of the relativistic electrons (Gitti et al. 2004, 2007a). However, there also seems to be some observational and theoretical evidence to support hadronic origin (Kempner et al. 2004 and references therein). Additionally, in two clusters (A 2142: Markevitch et al. 2000; RXJ 1347.4–1145: Gitti et al. 2007a, 2007b), we got indications that cluster mergers and cooling flows may act simultaneously in powering mini-halo emission in the rare and peculiar clusters in which they coexist. Further theoretical and observational studies are indeed essential due to the low number of known radio mini-halos (Sect. 2.2).

4 Measurement of Intracluster Magnetic Fields

As stressed above, the presence of diffuse and extended synchrotron emission in galaxy clusters indicates the existence of weak magnetic fields in the cluster volume. Different possibilities for their origin have been proposed which are reviewed by Dolag et al. (2008—Chap. 15, this issue). Radio observations of galaxy clusters allow us to measure intracluster magnetic fields and test the different theories on their origin, as reviewed by Carilli and Taylor (2002) and Govoni and Feretti (2004). In the following the main methods to study magnetic field intensity and, eventually, structure are summarised.

4.1 Equipartition Magnetic Fields

In the optically thin case, the total monochromatic emissivity $J(\nu)$ from a set of relativistic electrons in a magnetic field **B** depends on a) the magnetic field strength, b) the energy distribution of the electrons, which is usually assumed to be a power law (1), and c) the pitch angle between the electron velocity and the magnetic field direction (θ)

$$J(\nu) \propto N_0 (B \sin\theta)^{1+\alpha} \nu^{-\alpha}, \tag{3}$$

where $\alpha = (\delta - 1)/2$ is the spectral index of the synchrotron spectrum.[4]

Synchrotron emission from diffuse and extended radio sources can give us a direct measure for the intensity of cluster magnetic fields if the relativistic electron flux is measured or constrained. That can be achieved, for example, if Compton-produced X-ray (and gamma-ray) emission was detected simultaneously (see Sect. 4.2). In the case of polarised radio emission, we can also get an indication of the projected magnetic field orientation and its degree of ordering. To break the degeneracy between magnetic field strength and electron density (3), and to obtain a measure for cluster magnetic fields from the observed luminosity of radio sources, it is typically assumed that the energy density of the relativistic plasma within a radio source is minimum

$$U_{\rm tot} = U_{\rm el} + U_{\rm pr} + U_B = U_{\rm min}, \tag{4}$$

where U_B is the energy density in magnetic fields, and $U_{\rm el}$ and $U_{\rm pr}$ are the energy in electrons and in protons respectively. The energy in the heavy particles (protons) is considered to be related to the electron energy

$$U_{\rm pr} = k U_{\rm el}. \tag{5}$$

The value of k depends on the mechanism of (re-)acceleration of electrons, whose physical details, as seen above, are still unknown. A typical value of $k = 1$ is adopted for halo and relic sources. Another important assumption of this method relates to the value of the filling factor, Φ, i.e. the fraction of the source volume V occupied by magnetic field and relativistic particles. The energy density in magnetic field is given by

$$U_B = \frac{B^2}{8\pi} \Phi V. \tag{6}$$

[4] δ is the electron energy index, see (1).

It is usually considered that particles and magnetic fields occupy the entire volume, i.e. $\Phi = 1$. It can be derived easily that the condition of minimum energy is obtained when the contributions of cosmic rays and magnetic fields is approximately equal

$$U_B = \frac{3}{4}(1+k)U_{\rm el}. \tag{7}$$

This is the so-called classical equipartition assumption, which allows us to estimate the magnetic field of a radio source from its radio luminosity L (see Pacholczyk 1970 for a rigorous derivation)

$$B_{\rm eq} \propto \left[\frac{L(1+k)}{\Phi V}\right]^{2/7}. \tag{8}$$

In the standard approach presented above, L is the observed synchrotron luminosity between two fixed frequencies ν_1 and ν_2 (usually $\nu_1 = 10$ MHz and $\nu_2 = 100$ GHz). In this way, however, the integration limits are variable in terms of the energy of the radiating electrons, since, based on (3), electron energies corresponding to ν_1 and ν_2 depend on magnetic field values. This point is particularly relevant for the lower limit, owing to the power-law shape of the electron energy distribution and to the expected presence of low energy electrons in radio halos/relics. Alternatively, it has been suggested to derive equipartition quantities by integrating the electron luminosity over an energy range ($\gamma_{\rm min} - \gamma_{\rm max}$) (Brunetti et al. 1997; Beck and Krause 2005). It can be shown that, for $\gamma_{\rm min} \ll \gamma_{\rm max}$ and $\alpha > 0.5$, the new expression for the equipartition magnetic field is

$$B'_{\rm eq} \sim 1.1 \gamma_{\rm min}^{\frac{1-2\alpha}{3+\alpha}} B_{\rm eq}^{\frac{7}{2(3+\alpha)}}, \tag{9}$$

where $B_{\rm eq}$ is the equipartition magnetic field expressed in Gauss derived through (8). Typically, for $B_{\rm eq} \sim \mu$G, $\gamma_{\rm min} \sim 100$ and $\alpha \sim 0.75$–1, this new approach gives magnetic field values 2 to 5 times larger than the standard method.

Estimates of equipartition fields on scales as large as ∼1 Mpc give magnetic field intensities in the range 0.1–1 μG. As we have seen, these estimates are based on several assumptions both on different physical properties of the radio emitting region (e.g., the filling factor Φ and the ratio between electron and proton energies k), and on the condition of minimum energy of the observed relativistic plasma. Since the validity of these assumptions is not obvious, one has to be aware of the uncertainties and thus of the limits inherent to the equipartition determination of magnetic fields.

4.2 Compton Scattering of CMB Photons

As reviewed by Rephaeli et al. (2008—Chap. 5, this issue), 3K microwave background photons can be subject to Compton scattering by electrons in the cluster volume. If the presence of thermal particles in the ICM results in a distortion of the Cosmic Microwave Background (CMB) spectrum well known as "Sunyaev-Zel'dovich effect" (Sunyaev and Zel'dovich 1972), non-thermal hard X-ray (HXR) photons are produced via Compton scattering by the same cosmic rays that are responsible for the synchrotron emission observed at radio wavelengths. Compton scattering increases the frequency of the incoming photon through

$$\nu_{\rm out} = \frac{4}{3}\gamma^2 \nu_{\rm in}. \tag{10}$$

The Planck function of the CMB peaks at $\nu_{\rm in} \sim 1.6 \times 10^{11}$ Hz. Based on (10), for typical energies of relativistic electrons in clusters ($\gamma \sim 1000$–5000), the scattered photons fall in the X-ray and gamma-ray domain ($\sim 2 \times 10^{17}$–5×10^{18} Hz, i.e. $\sim$0.8–20.7 keV).

Non-thermal HXR emission from galaxy clusters due to Compton scattering of CMB photons was predicted more than 30 years ago (e.g., Rephaeli 1977) and has now been detected in several systems (Rephaeli et al. 2008—Chap. 5, this issue; Fusco-Femiano et al. 2007 and references therein). Alternative interpretations to explain the detected non-thermal X-ray emission have been proposed in the literature (Blasi and Colafrancesco 1999; Enßlin et al. 1999; Blasi 2000; Dogiel 2000; Sarazin and Kempner 2000). However, these hypotheses seem to be ruled out by energetic considerations, because of the well known inefficiency of the proposed non-thermal Bremsstrahlung (NTB) mechanism. NTB emission of keV regime photons with some power P immediately imply about 10^5 times larger power to be dissipated in the plasma that seems to be unrealistic in a quasi-steady model (Petrosian 2001, 2003). For a more detailed treatment of the origin of HXR emission from galaxy clusters, see the review by Petrosian et al. (2008—Chap. 10, this issue).

The detection of non-thermal HXR and radio emission, produced by the same population of relativistic electrons, allows us to estimate unambiguously the volume-averaged intracluster magnetic field. Following the exact derivations by Blumenthal and Gould (1970), the equations for the synchrotron flux $f_{\rm syn}$ at the frequency ν_R and the Compton X-ray flux $f_{\rm C}$ at the frequency ν_X are

$$f_{\rm syn}(\nu_R) = \frac{\Phi V}{4\pi {D_{\rm L}}^2} \frac{4\pi {\rm e}^3}{(m_{\rm e}{\rm c}^2)^{\delta}} N_0 B^{\frac{\delta+1}{2}} \left(\frac{3{\rm e}}{4\pi m_{\rm e}{\rm c}}\right)^{\frac{\delta-1}{2}} a(\delta) \nu_R^{-\frac{\delta-1}{2}}, \tag{11}$$

$$f_{\rm C}(\nu_X) = \frac{\Phi V}{4\pi {D_{\rm L}}^2} \frac{8\pi^2 {\rm r}_0^2}{{\rm c}^2} {\rm h}^{-\frac{\delta+3}{2}} N_0 (m_{\rm e}{\rm c}^2)^{(1-\delta)} ({\rm k}T)^{\frac{\delta+5}{2}} F(\delta) \nu_X^{-\frac{\delta-1}{2}}. \tag{12}$$

Here h is the Planck constant, V is the volume of the source and Φ the filling factor, $D_{\rm L}$ is the luminosity distance of the source, B the magnetic field strength, T the radiation temperature of the CMB, ${\rm r}_0$ the classical electron radius (or Thomson scattering length), N_0 and δ are the amplitude and the spectral index of the electron energy distribution (1). The values of the functions $a(\delta)$ and $F(\delta)$ for different values of δ can be found in Blumenthal and Gould (1970). The field B can thus be estimated directly from these equations

$$B \propto \left(\frac{f_{\rm syn}(\nu_R)}{f_{\rm C}(\nu_X)}\right)^{\frac{2}{\delta+1}} \left(\frac{\nu_R}{\nu_X}\right)^{\frac{\delta-1}{\delta+1}}. \tag{13}$$

Typical cluster magnetic field values of $\sim$0.1–0.3 μG are obtained (e.g., Rephaeli et al. 1999; Fusco-Femiano et al. 1999, 2000, 2001; Rephaeli and Gruber 2003; Rephaeli et al. 2006). Compared to equipartition measures, this method has the great advantage of using only observables, assuming only that the spatial factors in the expressions for the synchrotron and Compton fluxes (Rephaeli 1979) are identical.

4.3 Faraday Rotation Measure

Faraday rotation analysis of radio sources in the background or in the galaxy clusters themselves is one of the key techniques used to obtain information on the cluster magnetic fields. The presence of a magnetised plasma between an observer and a radio source changes the properties of the polarised emission from the radio source. Therefore information on cluster

magnetic fields along the line-of-sight can be determined, in conjunction with X-ray observations of the hot gas, through the analysis of the Rotation Measure (RM) of radio sources (e.g., Burn 1966).

The polarised synchrotron radiation coming from radio galaxies undergoes the following rotation of the plane of polarisation as it passes through the magnetised and ionised intracluster medium

$$\Psi_{\rm Obs}(\lambda) = \Psi_{\rm Int} + \lambda^2 \times RM, \tag{14}$$

where $\Psi_{\rm Int}$ is the intrinsic polarisation angle, and $\Psi_{\rm Obs}(\lambda)$ is the polarisation angle observed at a wavelength λ. The RM is related to the thermal electron density ($n_{\rm e}$), the magnetic field along the line-of-sight ($B_{\parallel}$), and the path-length (L) through the intracluster medium according to

$$RM_{[\rm rad\,m^{-2}]} = 812 \int_0^{L_{[\rm kpc]}} n_{\rm e\,[cm^{-3}]} B_{\parallel\,[\mu\rm G]}\,{\rm d}l. \tag{15}$$

Polarised radio galaxies can be mapped at several frequencies to produce, by fitting (14), detailed RM images. Once the contribution of our Galaxy is subtracted, the RM should be dominated by the contribution of the intracluster medium, and therefore it can be used to estimate the cluster magnetic field strength along the line of sight.

The RM observed in radio galaxies may not be all due to the cluster magnetic field if the RM gets locally enhanced by the intracluster medium compression due to the motion of the radio galaxy itself. However a statistical RM investigation of point sources (Clarke et al. 2001; Clarke 2004) shows a clear broadening of the RM distribution toward small projected distances from the cluster centre, indicating that most of the RM contribution comes from the intracluster medium. This study included background sources, which showed similar enhancements as the embedded sources.

We also note that there are inherent uncertainties in the determination of field values from Faraday Rotation measurements, stemming largely from the unknown small-scale tangled morphology of intracluster fields, their large-scale spatial variation across the cluster, and from the uncertainty in modelling the gas density profile (see, e.g., Goldshmidt and Rephaeli 1993; Newman et al. 2002; Rudnick and Blundell 2003a; Enßlin et al. 2003; Murgia et al. 2004).

RM studies of radio galaxies have been carried out on both statistical samples (e.g., Lawler and Dennison 1982; Vallée et al. 1986; Kim et al. 1990, 1991; Clarke et al. 2001) and individual clusters by analysing detailed high resolution RM images (e.g., Perley and Taylor 1991; Taylor and Perley 1993; Feretti et al. 1995, 1999; Govoni et al. 2001a; Taylor et al. 2001; Eilek and Owen 2002; Govoni et al. 2006; Taylor et al. 2007; Guidetti et al. 2008). Both for interacting and relaxed (cooling flow) clusters the RM distribution of radio galaxies is generally patchy, indicating that cluster magnetic fields have structures on scales as low as 10 kpc or less. RM data are usually consistent with central magnetic field strengths of a few μG. But, radio galaxies at the centre of relaxed clusters have extreme RM, with the magnitude of the RM roughly proportional to the cooling flow rate (see Fig. 6). Strong magnetic fields are derived in the high density cooling-core regions of some clusters, with values exceeding ∼10 μG (e.g., in the inner region of Hydra A, a value of ∼35 μG was deduced by Taylor et al. 2002). It should be emphasised that such high field values are clearly not representative of the mean fields in large extended regions.

Dolag et al. (2001b) showed that, in the framework of hierarchical cluster formation, the correlation between two observable parameters, the RM and the cluster X-ray surface brightness in the source location, is expected to reflect a correlation between the cluster magnetic

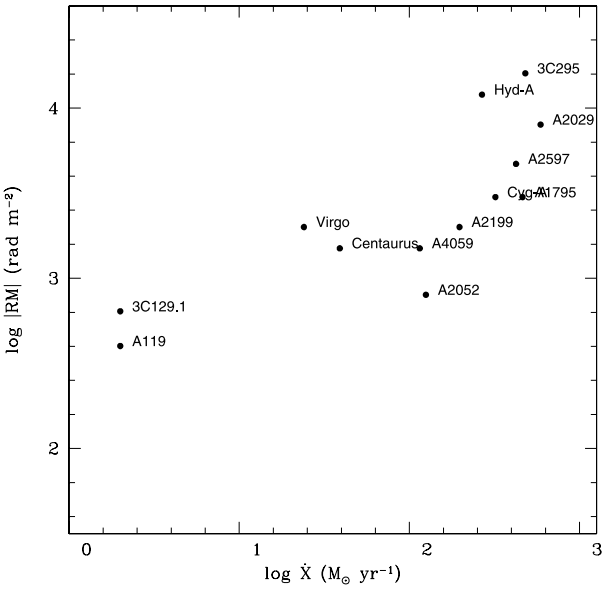


Fig. 6 RM magnitudes of a sample of radio galaxies located in cooling flow clusters, plotted as a function of the cooling flow rate $\dot{X}$ (from Taylor et al. 2002)

field and gas density. Therefore, from the analysis of the RM versus X-ray brightness it is possible to infer the trend of magnetic field versus gas density.

On the basis of the available high quality RM images, increasing attention is given to the power spectrum of the intracluster magnetic field fluctuations. Several studies (Enßlin and Vogt 2003; Murgia et al. 2004) have shown that detailed RM images of radio galaxies can be used to infer not only the cluster magnetic field strength, but also the cluster magnetic field power spectrum. The analyses of Vogt and Enßlin (2003, 2005) and Guidetti et al. (2008) suggest that the power spectrum is of the Kolmogorov type, if the auto-correlation length of the magnetic fluctuations is of the order of few kpc. However, Murgia et al. (2004) and Govoni et al. (2006) pointed out that shallower magnetic field power spectra are possible if the magnetic field fluctuations extend out to several tens of kpc.

4.4 Comparison of the Different Methods

As shown in Table 3 of Govoni and Feretti (2004), the different methods available to measure intracluster magnetic fields show quite discrepant results (even more than a factor 10). RM estimates are about an order of magnitude higher than the measures derived both from the synchrotron diffuse radio emission and the non-thermal hard X-ray emission ($\sim$1–5 μG vs. $\sim$0.2–1 μG).

This can be due to several factors. Firstly, equipartition values are severely affected by the already mentioned physical assumptions of this method. Secondly, while RM estimates give a weighted average of the field along the line of sight, equipartition and Compton scattering measures are made by averaging over larger volumes. Additionally, discrepancies can be due to spatial profiles of both the magnetic field and the gas density not being constant all over the cluster (Goldshmidt and Rephaeli 1993), or due to compressions, fluctuations and inhomogeneities in the gas and in the magnetic field, related to the presence of radio galaxies or to the dynamical history of the cluster (e.g., on-going merging events) (Beck et al. 2003; Rudnick and Blundell 2003b; Johnston-Hollitt 2004). Finally, a proper modelling of the Compton scattering method should include a) the effects of aged electron spectra,

b) the expected radial profile of the magnetic field, and c) possible anisotropies in the pitch angle distribution of electrons (Brunetti et al. 2001; Petrosian 2001).

An additional method of estimating cluster magnetic fields comes from the X-ray analysis of cold fronts (Vikhlinin et al. 2001). These X-ray cluster features, discovered by Markevitch et al. (2000) thanks to the exquisite spatial resolution of the *Chandra* satellite, result from dense cool gas moving with near-sonic velocities through the less dense and hotter ICM. Cold fronts are thus subject to Kelvin-Helmholtz (K-H) instability that, for typical cluster and cold front properties (Mach number, gas temperatures, cluster-scale length), could quickly disturb the front outside a narrow ($\lesssim 10°$) sector in the direction of the cool cloud motion.[5] Through the *Chandra* observation of A 3667, Vikhlinin et al. (2001) instead revealed a cold front that is stable within a $\pm 30°$ sector. They showed that a ~ 10 μG magnetic field oriented nearly parallel to the front is able to suppress K-H instability, thus preserving the front structure, in a $\pm 30°$ sector. The estimated magnetic field value, significantly higher than the typical measures given by the other methods outside cluster cooling flows, is likely an upper limit of the absolute field strength. Near the cold front the field is actually amplified by tangential gas motions (see Vikhlinin et al. 2001).

Variations of the magnetic field structure and strength with the cluster radius have been recently pointed out by Govoni et al. (2006). By combining detailed multi-wavelength and numerical studies we will get more insight into the strength and structure of intracluster magnetic fields, and into their connection with the thermodynamical evolution of galaxy clusters. More detailed comparisons of the different approaches for measuring intracluster magnetic fields can be found, for instance, in Petrosian (2003) and Govoni and Feretti (2004).

5 Diffuse Radio Emission in Galaxy Clusters: Open Questions and Perspectives

Significant progress has been made recently in our knowledge on the non-thermal component of galaxy clusters. A number of open questions arise in assessing the current theoretical and observational status.

First of all, we need to test the current theories on the origin of the large-scale non-thermal component in clusters (magnetic field and cosmic rays). If at present primary models seem to be the favourite acceleration mechanisms for intracluster electrons, secondary models cannot be ruled out. Among other things, it will be necessary to establish: How common is the non-thermal component in clusters? Is it really hosted *only* in merging systems (as present observational results suggest) or do *all* clusters have a radio halo/relic? If this latter hypothesis is correct, how should we modify the radio power versus X-ray luminosity correlation (Sect. 2.1)? If shocks and turbulence related to cluster mergers are instead the mechanisms responsible for electron re-acceleration, why have extended radio sources not been detected in *all* merging clusters? Is this related to other physical effects (i.e. the merging event alone is not enough to produce intracluster cosmic rays), as the correlation between radio power and cluster mass seems to indicate, or it is due to a lack of sensitivity of the current instruments (i.e. all merging clusters host radio halos and relics, but a large fraction of these sources lies below the sensitivity limit of present telescopes)?

Among the previous questions, the most difficult to answer *at present* are those that involve the study of low-luminous X-ray clusters, for which the limits of current radio

[5] For a more precise treatment, see Vikhlinin et al. (2001) and Markevitch and Vikhlinin (2007).

observations are particularly severe. By extrapolating to low radio and X-ray luminosity the P_ν–L_X relation (Sect. 2.1), Feretti and Giovannini (2008) have estimated that, if present, halos with typical sizes of 1 Mpc in intermediate/low-luminous X-ray clusters ($L_{X[0.1-2.4\ \mathrm{keV}]} \lesssim 5 \times 10^{44}\ \mathrm{h}_{70}^{-2}\ \mathrm{erg\ s}^{-1}$)[6] would actually have a radio surface brightness lower than the current limits obtained in the literature and in the NVSS. At higher X-ray luminosities, more constraints on radio halo statistics have recently been obtained by Venturi et al. (2007) and Brunetti et al. (2007). They carried out GMRT[7] observations at 610 MHz of 34 luminous ($L_{X[0.1-2.4\ \mathrm{keV}]} \gtrsim 5 \times 10^{44}\ \mathrm{h}_{70}^{-2}\ \mathrm{erg\ s}^{-1}$) clusters with $0.2 \lesssim z \lesssim 0.4$. The bulk of the galaxy clusters in their sample does not show any diffuse central radio emission, with radio luminosity upper limits that are well below the P_ν–L_X relation derived from the previously known radio halos. The net bimodality of the cluster distribution in the P_ν–L_X plane support primary models against secondary models. Actually, the former predict a relatively fast ($\approx 10^8$ yrs) transition of clusters from a radio quiet state to the observed P_ν–L_X correlation, where they remain for $\lesssim$1 Gyr. A significantly wider scatter around the P_ν–L_X correlation is instead expected in the frame of secondary models, that could be reconciled with observations only assuming the existence of strong dissipation of magnetic fields in clusters (see Brunetti et al. 2007, and references therein).

On smaller scales, a larger sample of radio mini-halos is required to test the current theories on the origin of radio emission in this class of sources (Gitti et al. 2002; Pfrommer and Enßlin 2004). Recent results suggest that cooling flows and mergers could act simultaneously, when they co-exist, in providing energy to the relic population of relativistic electrons injected into the ICM by AGNs, thus powering mini-halo radio emission (Gitti et al. 2007a).

As discussed in Sect. 4, radio observations of galaxy clusters offer a unique tool to estimate strength and structure of large-scale magnetic fields, allowing to test the different scenarios of their origin. Several observational results show that magnetic fields of the order of $\sim\mu$G are common in clusters. Through combined numerical and observational analyses, Murgia et al. (2004) and Govoni et al. (2006) have shown that detailed morphology and polarisation information of radio halos may provide important constraints on the strength and structure of intracluster magnetic fields. However, discrepant results have been obtained up to now (Sect. 4.4) and more detailed information on magnetic fields is still needed.

A better knowledge of the physics of the non-thermal component in galaxy clusters will have important cosmological implications. If it will be confirmed that the presence of giant halos and relics is related to cluster mergers, the statistical properties of these radio sources will allow us to test the current cluster formation scenario, giving important hints on large-scale structure formation and, thus, cosmological parameters (e.g., Evrard and Gioia 2002).

Additionally, we will be able to estimate how the gravitational energy released during cluster mergers is redistributed between the thermal ICM and the relativistic plasma (e.g., Sarazin 2005). The effects of magnetic fields on the thermodynamical evolution of large-scale structures will be evaluated, as well as the contribution of the non-thermal pressure to the estimate of mass and temperature in galaxy clusters (e.g., Dolag and Schindler 2000; Dolag et al. 2001a; Colafrancesco et al. 2004). Cluster scaling laws, such as mass vs. temperature, are actually key ingredients to derive cosmological constraints from galaxy clusters (e.g., Ettori et al. 2004; Arnaud et al. 2005).

[6] Converted from the bolometric X-ray luminosity limit in Feretti and Giovannini (2008) to the 0.1–2.4 keV band luminosity using Table 5 of Böhringer et al. (2004) and assuming typical ICM temperature values ($T_X \sim$ 5–10 keV).

[7] The Giant Metrewave Radio Telescope (GMRT) is operated by the National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research (NCRA-TIFR).

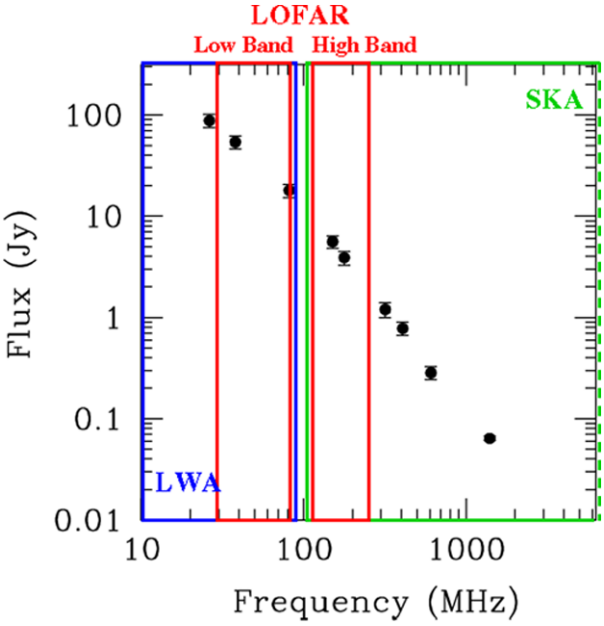


Fig. 7 Spectrum of the diffuse radio source in A 1914 (from Bacchi et al. 2003). Superimposed the frequency range covered by LWA (10–88 MHz, *in blue*), LOFAR (Low Band: 30–80 MHz, High Band: 110–240 MHz, *in red*) and SKA (100–25 GHz, *in green*). The low-frequency domain covered by the next generation radio-telescopes is optimal for the detection of high spectral index radio sources, such as radio halos, mini-halos and relics

Finally, a better knowledge of extended radio sources in clusters is indeed essential for complementary cosmological studies, e.g., the epoch of re-ionisation (EoR). It has been proven that radio halos and relics are the strongest extra-galactic foreground sources to be removed in order to probe the EoR through the study of the redshifted 21 cm emission from neutral hydrogen (e.g., Di Matteo et al. 2004). Better models for the diffuse radio emission have to be inserted into numerical simulations of the EoR 21 cm emission, in order to understand how to remove efficiently the contamination due to radio halos and relics.

An increase in the number of known radio halos/relics, as well as higher resolution and sensitivity observations, are essential to answer the main open questions, summarised at the beginning of this section, about the nature of diffuse radio emission in clusters. As shown in Fig. 7, halos and relics are difficult to detect in the GHz range due to their steep spectra. Several observations performed with the currently available low-frequency instruments (e.g., GMRT: Venturi et al. 2007; VLA: Kassim et al. 2001; Orrù et al. 2007) confirm the interest in studying this class of radio sources at high wavelengths. A short term perspective in the study of radio halos and relics is thus to fully exploit those instruments that are already available for observations in the MHz range of the electromagnetic spectrum with good enough sensitivity (approximately from some tens of μJy/beam to some mJy/beam) and angular resolution (roughly some tens of arcsec). However, in order to make a proper comparison between observational results and current theoretical models about the origin of radio halos and relics, we need multi-frequency observations of *statistical* samples of diffuse radio sources. Current telescopes require too long exposure-time per cluster ($\gtrsim$1–2 hours) to reach the sensitivity limits necessary for detecting radio halos/relics, making statistical analyses of diffuse radio emission in clusters extremely time-demanding.

The low-frequency range covered by a new generation of radio telescopes (Long Wavelength Array—LWA; Low Frequency Array—LOFAR; Square Kilometre Array—SKA), together with their gain in sensitivity and resolution, will increase dramatically the statistics on the number of known radio halos and relics. Not only these instruments will cover the

optimal frequency range for halo/relic detection (see Fig. 7), but also their gain in sensitivity and resolution will be of the order of 10 to 1000 (see Table 2 of Brüggen et al. 2005), allowing observations of statistical samples of diffuse and extended radio sources. A LOFAR survey at 120 MHz, covering half the sky to a 5 σ flux limit of 0.1 mJy (1 hour integration time per pointing), could detect $\sim$1000 halos/relics, of which 25% at redshift larger than $z \sim 0.3$ (Röttgering 2003). Feretti et al. (2004b) have estimated that, with 1 hour integration time at 1.4 GHz, 50% of the SKA collecting area will allow us to detect halos and relics of total flux down to 1 mJy at any redshift, and down to 0.1 mJy at high redshift. Based on our current knowledge, more than three (fifteen) hundred diffuse cluster radio sources are expected at 1.5 GHz on the full sky at the 1 mJy (0.1 mJy) flux limit (Enßlin and Röttgering 2002).

With statistical samples of halos and relics over a wide redshift range, we will be able to a) test the correlation between the non-thermal component and the physical properties of clusters (dynamical state, mass, X-ray luminosity and temperature...), b) analyse the redshift evolution of halos and relics, with the advantages for cosmological studies stressed above, and c) fill the gap in our knowledge of the last phases of radio galaxy evolution in clusters (see Sect. 2.3). Particularly interesting will be the study of possible presence of non-thermal radio emission at $z \sim 1$, i.e. the epoch of formation of the massive galaxy clusters observed in the local Universe.

The excellent sensitivity, high angular resolution and large number of spectral channels of the next generation instruments, together with new techniques of RM synthesis (Brentjens and de Bruyn 2005), will allow polarisation mapping and RM studies of radio emission in clusters, significantly improving our estimates of large-scale magnetic fields.

Future radio observations of galaxy clusters, combined with the new generation instruments at other wavelengths[8] (e.g., sub-mm: ALMA; X-ray: XEUS, Simbol-X; gamma-rays: GLAST, H.E.S.S., MAGIC;[9]...), will allow us to open a new window in cosmological studies.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". CF and FG warmly thank Luigina Feretti and Matteo Murgia for many useful discussions on the subject of this paper. CF and SS acknowledge financial support by the Austrian Science Foundation (FWF) through grants P18523-N16 and P19300-N16, by the Tiroler Wissenschaftsfonds and through the UniInfrastrukturprogramm 2005/06 by the BMWF. FG acknowledges financial support through Grant ASI-INAF I/088/06/0—High Energy Astrophysics.

References

S.W. Allen, S. Ettori, A.C. Fabian, Mon. Not. R. Astron. Soc. **324**, 877 (2001)
M. Arnaud, E. Pointecouteau, G.W. Pratt, Astron. Astrophys. **441**, 893 (2005)
M. Bacchi, L. Feretti, G. Giovannini, F. Govoni, Astron. Astrophys. **400**, 465 (2003)
J. Bagchi, T. Enßlin, F. Miniati et al., New Astron. **7**, 249 (2002)
J. Bagchi, F. Durret, G.B. Lima Neto, S. Paul, Science **314**, 791 (2006)
S.A. Baum, C.P. O'Dea, Mon. Not. R. Astron. Soc. **250**, 737 (1991)
R. Beck, M. Krause, Astron. Nachr. **326**, 414 (2005)
R. Beck, D. Sukurov, A., Sokoloff, R. Wielebinski, Astron. Astrophys. **411**, 99 (2003)

[8] See also Paerels et al. (2008—Chap. 19, this issue).

[9] A new generation of ground based imaging Cherenkov telescopes will be available soon. The measurements above 10 TeV are crucial to distinguish between the Compton scattering and hadronic origins of gamma-ray emission from clusters of galaxies.

V.S. Berezinsky, P. Blasi, V.S. Ptuskin, Astrophys. J. **487**, 529 (1997)
P. Blasi, Astrophys. J. **532**, L9 (2000)
P. Blasi, J. Korean Astron. Soc. **37**, 483 (2004)
P. Blasi, S. Colafrancesco, Astrophys. Phys. **12**, 169 (1999)
P. Blasi, S. Gabici, G. Brunetti, Int. J. Mod. Phys. A **22**, 681 (2007)
G.R. Blumenthal, R.J. Gould, Rev. Mod. Phys. **42**, 237 (1970)
H. Böhringer, P. Schuecker, L. Guzzo et al., Astron. Astrophys. **425**, 367 (2004)
M.A. Brentjens, A.G. de Bruyn, Astron. Astrophys. **441**, 1217 (2005)
M. Brüggen, H. Falcke, R. Beck et al., German LOFAR white paper, 2005. Available electronically at http://lofar.mpa-garching.mpg.de/glow.html
G. Brunetti, J. Korean Astron. Soc. **37**, 493 (2004)
G. Brunetti, G. Setti, A. Comastri, Astron. Astrophys. **325**, 898 (1997)
G. Brunetti, G. Setti, L. Feretti, G. Giovannini, Mon. Not. R. Astron. Soc. **320**, 365 (2001)
G. Brunetti, P. Blasi, R. Cassano, S. Gabici, Mon. Not. R. Astron. Soc. **350**, 1174 (2004)
G. Brunetti, T. Venturi, D. Dallacasa et al., Astrophys. J. **670**, L5 (2007)
D.A. Buote, Astrophys. J. **553**, L15 (2001)
D.A. Buote, in *Merging Processes in Galaxy Clusters*, ed. by L. Feretti, I.M. Gioia, G. Giovannini. Astrophys. Sp. Sc. Lib. (Kluwer), **272**, 79 (2002)
D.A. Buote, J.C. Tsai, Astrophys. J. **458**, 27 (1996)
B.J. Burn, Mon. Not. R. Astron. Soc. **133**, 67 (1966)
J.O. Burns, M.E. Sulkanen, G.R. Gisler, R.A. Perley, Astrophys. J. **388**, L49 (1992)
A.M. Bykov, H. Bloemen, Yu.A. Uvarov, Astron. Astrophys. **362**, 886 (2000)
A.M. Bykov, K. Dolag, F. Durret, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9312-9
C.L. Carilli, G.B. Taylor, Annu. Rev. Astron. Astrophys. **40**, 319 (2002)
R. Cassano, G. Brunetti, Mon. Not. R. Astron. Soc. **357**, 1313 (2005)
R. Cassano, G. Brunetti, G. Setti, Mon. Not. R. Astron. Soc. **369**, 1577 (2006)
R. Cassano, G. Brunetti, G. Setti, F. Govoni, K. Dolag, Mon. Not. R. Astron. Soc. **378**, 1565 (2007)
T.E. Clarke, J. Korean Astron. Soc. **37**, 337 (2004)
T.E. Clarke, T.A. Enßlin, Astron. J. **131**, 2900 (2006)
T.E. Clarke, P.P. Kronberg, H. Böhringer, Astrophys. J. **547**, L111 (2001)
S. Colafrancesco, in *Diffuse Tthermal and Relativistic Plasma in Galaxy Clusters*, ed. by H. Böhringer, L. Feretti, P. Schuecker, MPE Rep. **271**, 269 (1999). astro-ph/9907329
S. Colafrancesco, A. Dar, A. De Rújula, Astron. Astrophys. **413**, 441 (2004)
J.J. Condon, W.D. Cotton, E.W. Greisen et al., Astron. J. **115**, 1693 (1998)
B.M. Deiss, W. Reich, H. Lesch, R. Wielebinski, Astron. Astrophys. **321**, 55 (1997)
K.M. Delain, L. Rudnick, Astron. Nachr. **327**, 561 (2006)
B. Dennison, Astrophys. J. **239**, L93 (1980)
T. Di Matteo, B. Ciardi, F. Miniati, Mon. Not. R. Astron. Soc. **355**, 1053 (2004)
V.A. Dogiel, Astron. Astrophys. **357**, 66 (2000)
K. Dolag, T.A. Enßlin, Astron. Astrophys. **362**, 151 (2000)
K. Dolag, S. Schindler, Astron. Astrophys. **364**, 491 (2000)
K. Dolag, A. Evrard, M. Bartelmann, Astron. Astrophys. **369**, 36 (2001a)
K. Dolag, S. Schindler, F. Govoni, L. Feretti, Astron. Astrophys. **378**, 777 (2001b)
K. Dolag, A.M. Bykov, A. Diaferio, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9319-2
J.A. Eilek, F.N. Owen, Astrophys. J. **567**, 202 (2002)
T.A. Enßlin, Gopal-Krishna, Astron. Astrophys. **366**, 26 (2001)
T.A. Enßlin, H.J.A. Röttgering, Astron. Astrophys. **396**, 83 (2002)
T.A. Enßlin, C. Vogt, Astron. Astrophys. **401**, 835 (2003)
T.A. Enßlin, P.L. Biermann, U. Klein, S. Kohle, Astron. Astrophys. **332**, 395 (1998)
T.A. Enßlin, R. Lieu, P.L. Biermann, Astron. Astrophys. **344**, 409 (1999)
T.A. Enßlin, C. Vogt, T.E. Clarke, G.B. Taylor, Astrophys. J. **597**, 870 (2003)
S. Ettori, A.C. Fabian, D.A. White, Mon. Not. R. Astron. Soc. **300**, 837 (1998)
S. Ettori, P. Tozzi, S. Borgani, P. Rosati, Astron. Astrophys. **417**, 13 (2004)
A.E. Evrard, I.M. Gioia, in *Merging Processes in Galaxy Clusters*, ed. by L. Feretti, I.M. Gioia, G. Giovannini, Astrophys. Sp. Sc. Lib. (Kluwer), **272**, 253 (2002)
A.C. Fabian, J.S. Sanders, S. Ettori et al., Mon. Not. R. Astron. Soc. **318**, 65 (2000)
A.C. Fabian, A. Celotti, K.M. Blundell, N.E. Kassim, R.A. Perley, Mon. Not. R. Astron. Soc. **331**, 369 (2002)
L. Feretti, in *Diffuse Thermal and Relativistic Plasma in Galaxy Clusters*, ed. by H. Böhringer, L. Feretti, P. Schuecker, MPE Rep. **271**, 3 (1999)
L. Feretti, in *The Universe at Low Radio Frequencies*, ed. by Pramesh Rao, G. Swarup, Gopal-Krishna. IAU Symp. **199**, 133 (2002)

L. Feretti, Adv. Space Res. **36**, 729 (2005)
L. Feretti, G. Giovannini, in *Panchromatic View of Clusters of Galaxies and the Large-Scale Structure*, ed. by M. Plionis, O. Lopez-Cruz, D. Hughes (2008). astro-ph/0703494
L. Feretti, D.M. Neumann, Astron. Astrophys. **450**, L21 (2006)
L. Feretti, T. Venturi, in *Merging Processes in Galaxy Clusters*, ed. by L. Feretti, I.M. Gioia, G. Giovannini, Astrophys. Sp. Sc. Lib. (Kluwer) **272**, 163 (2002)
L. Feretti, D. Dallacasa, G. Giovannini, A. Tagliani, Astron. Astrophys. **302**, 680 (1995)
L. Feretti, G. Giovannini, H. Böhringer, New Astron. **2**, 501 (1997)
L. Feretti, D. Dallacasa, F. Govoni et al., Astron. Astrophys. **344**, 472 (1999)
L. Feretti, R. Fusco-Femiano, G. Giovannini, F. Govoni, Astron. Astrophys. **373**, 106 (2001)
L. Feretti, E. Orrù, G. Brunetti et al., Astron. Astrophys. **423**, 111 (2004a)
L. Feretti, C. Burigana, T.A. Enßlin, in *Science with the Square Kilometre Array*, ed. by C. Carilli, S. Rawlings, New Astron. Rev. **48**, 1137 (2004b)
C. Ferrari, Multi-wavelength analysis of merging galaxy clusters, PhD Thesis, Nice (2003). Available electronically at http://tel.archives-ouvertes.fr/docs/00/04/85/55/PDF/tel-00010416.pdf
C. Ferrari, S. Maurogordato, A. Cappi, C. Benoist, Astron. Astrophys. **399**, 813 (2003)
C. Ferrari, C. Benoist, S. Maurogordato, A. Cappi, E. Slezak, Astron. Astrophys. **430**, 19 (2005)
C. Ferrari, M. Arnaud, S. Ettori, S. Maurogordato, J. Rho, Astron. Astrophys. **446**, 417 (2006a)
C. Ferrari, R.W. Hunstead, L. Feretti, S. Maurogordato, S. Schindler, Astron. Astrophys. **457**, 21 (2006b)
A. Finoguenov, H. Böhringer, Y.-Y. Zhang, Astron. Astrophys. **442**, 827 (2005)
T. Furusho, N.Y. Yamasaki, T. Ohashi, R. Shibata, H. Ezawa, Astrophys. J. **561**, L165 (2001)
R. Fusco-Femiano, D. Dal Fiume, L. Feretti et al., Astrophys. J. **513**, L21 (1999)
R. Fusco-Femiano, D. Dal Fiume, S. De Grandi et al., Astrophys. J. **534**, L7 (2000)
R. Fusco-Femiano, D. Dal Fiume, M. Orlandini et al., Astrophys. J. **552**, L97 (2001)
R. Fusco-Femiano, R. Landi, M. Orlandini, Astrophys. J. **654**, 9 (2007)
S. Gabici, P. Blasi, Astrophys. J. **583**, 695 (2003)
G. Gavazzi, L. Cortese, A. Boselli et al., Astrophys. J. **597**, 210 (2003)
S. Giacintucci, T. Venturi, G. Brunetti et al., Astron. Astrophys. **440**, 867 (2005)
S. Giacintucci, T. Venturi, S. Bardelli et al., New Astron. **11**, 437 (2006)
G. Giovannini, L. Feretti, New Astron. **5**, 335 (2000)
G. Giovannini, L. Feretti, in *Merging Processes in Galaxy Clusters*, ed. by L. Feretti, I.M. Gioia, G. Giovannini, Astrophys. Sp. Sc. Lib. (Kluwer) **272**, 197 (2002)
G. Giovannini, L. Feretti, T. Venturi, K.-T. Kim, P.P. Kronberg, Astrophys. J. **406**, 399 (1993)
G. Giovannini, M. Tordi, L. Feretti, New Astron. **4**, 141 (1999)
M. Girardi, A. Biviano, in *Merging Processes in Galaxy Clusters*, ed. by L. Feretti, I.M. Gioia, G. Giovannini, Astrophys. Sp. Sc. Lib. (Kluwer) **272**, 39 (2002)
M. Gitti, G. Brunetti, G. Setti, Astron. Astrophys. **386**, 456 (2002)
M. Gitti, G. Brunetti, L. Feretti, G. Setti, Astron. Astrophys. **417**, 1 (2004)
M. Gitti, C. Ferrari, W. Domainko, L. Feretti, S. Schindler, Astron. Astrophys. **470**, L25 (2007a)
M. Gitti, R. Piffaretti, S. Schindler, Astron. Astrophys. **472**, 383 (2007b)
O. Goldshmidt, Y. Rephaeli, Astrophys. J. **411**, 518 (1993)
P.L. Gómez, C. Loken, K. Röttiger, J.O. Burns, Astrophys. J. **569**, 122 (2002)
F. Govoni, L. Feretti, Int. J. Modern Phys. D **13**, 1549 (2004)
F. Govoni, G.B. Taylor, D. Dallacasa et al., Astron. Astrophys. **369**, 441 (2001a)
F. Govoni, L. Feretti, G. Giovannini et al., Astron. Astrophys. **376**, 803 (2001b)
F. Govoni, T.A. Enßlin, L. Feretti, G. Giovannini, Astron. Astrophys. **379**, 807 (2001c)
F. Govoni, M. Murgia, L. Feretti et al., Astron. Astrophys. **430**, L5 (2005)
F. Govoni, M. Murgia, L. Ferettiet al., Astron. Astrophys. **460**, 425 (2006)
D. Guidetti, M. Murgia, F. Govoni et al., Astron. Astrophys. (2008). arXiv:0709.2652
M. Hoeft, M. Brüggen, Mon. Not. R. Astron. Soc. **375**, 77 (2007)
W.J. Jaffe, Astrophys. J. **212**, 1 (1977)
M. Johnston-Hollitt, in *The Riddle of Cooling Flows in Galaxies and Clusters of Galaxies*, ed. by T. Reiprich, J. Kempner, N. Soker (2004). Published electronically at http://www.astro.virginia.edu/coolflow/
M. Johnston-Hollitt, R.W. Clay, R.D. Ekers, M.H. Wieringa, R.W. Hunstead, in *The Universe at low radio frequencies*, ed. by Pramesh Rao, G. Swarup, Gopal-Krishna. IAU Symp. **199**, 157 (2002)
C. Jones, W. Forman, Astrophys. J. **511**, 65 (1999)
W. Kapferer, C. Ferrari, W. Domainkoet al., Astron. Astrophys. **447**, 827 (2006)
N.E. Kassim, T.E. Clarke, T.A. Enßlin, A.S. Cohen, D.M. Neumann, Astrophys. J. **559**, 785 (2001)
J.C. Kempner, L.P. David, Mon. Not. R. Astron. Soc. **349**, 385 (2004)
J.C. Kempner, C.L. Sarazin, Astrophys. J. **548**, 639 (2001)

J.C. Kempner, E.L. Blanton, T.E. Clarke et al., in *The Riddle of Cooling Flows in Galaxies and Clusters of Galaxies*, ed. by T. Reiprich, J. Kempner, N. Soker (2004). Published electronically at http://www.astro.virginia.edu/coolflow/
K.T. Kim, P.P. Kronberg, P.E. Dewdney, T.L. Landecker, Astrophys. J. **355**, 29 (1990)
K.T. Kim, P.C. Tribble, P.P. Kronberg, Astrophys. J. **379**, 80 (1991)
M.I. Large, D.S. Mathewson, C.G. Haslam, Nature **183**, 1663 (1959)
J.M. Lawler, B. Dennison, Astrophys. J. **252**, 81 (1982)
H. Liang, R.W. Hunstead, M. Birkinshaw, P. Andreani, Astrophys. J. **544**, 686 (2000)
M.S. Longair, *High Energy Astrophysics* (Cambridge Univ. Press, Cambridge, 1981)
M. Markevitch, A. Vikhlinin, Astrophys. J. **563**, 95 (2001)
M. Markevitch, A. Vikhlinin, Phys. Rep. **443**, 1 (2007)
M. Markevitch, C.L. Sarazin, A. Vikhlinin, Astrophys. J. **521**, 526 (1999)
M. Markevitch, T.J. Ponman, P.E.J. Nulsen et al., Astrophys. J. **541**, 542 (2000)
M. Markevitch, F. Govoni, G. Brunetti, D. Jerius, Astrophys. J. **627**, 733 (2005)
S. Maurogordato, A. Cappi, C. Ferrari et al., Astron. Astrophys. (2008). astro-ph/0712.2715
B.R. McNamara, L. Bîrzan, D.A. Rafferty et al., in *Extragalactic Jets: Theory and Observation from Radio to Gamma Ray*, ed. by T.A. Rector, D.S. De Young, ASP Conference Series (2008). astro-ph/0708.0579
K. Meisenheimer, H.-J. Roser, P.R. Hiltner et al., Astron. Astrophys. **219**, 63 (1989)
F. Miniati, Mon. Not. R. Astron. Soc. **342**, 1009 (2003)
J. Mohr, M.J. Geller, G. Wegner, Astron. J. **112**, 1816 (1996)
J.S. Mulchaey, D.S. Davis, R.F. Mushotzky, D. Burstein, Astrophys. J. **456**, 80 (1996)
M. Murgia, F. Govoni, L. Feretti et al., Astron. Astrophys. **424**, 429 (2004)
M. Murgia, P. Parma, H.R. de Ruiter et al., in *X-Ray and Radio Connections*, ed. by L.O. Sjouwerman, K.K. Dyer (2005). Published electronically by NRAO, http://www.aoc.nrao.edu/events/xraydio
K. Nakazawa, K. Makishima, Y. Fukazawa, Publ. Astron. Soc. Jpn. **59**, 167 (2007)
D.M. Neumann, M. Arnaud, Astron. Astrophys. **348**, 711 (1999)
D.M. Neumann, M. Arnaud, Astron. Astrophys. **373**, L33 (2001)
W.I. Newman, A.I. Newman, Y. Rephaeli, Astrophys. J. **575**, 755 (2002)
E. Orrù, M. Murgia, L. Feretti et al., Astron. Astrophys. **467**, 943 (2007)
A.G. Pacholczyk, *Radio Astrophysics* (Freeman, New York, 1970)
F. Paerels, J. Kaastra, T. Ohashi et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9323-6
P. Parma, M. Murgia, H.R. de Ruiter et al., Astron. Astrophys. **470**, 875 (2007)
R.A. Perley, G.B. Taylor, Astron. J. **101**, 1623 (1991)
V. Petrosian, Astrophys. J. **557**, 560 (2001)
V. Petrosian, ASP Conf. Series **301**, 337 (2003)
V. Petrosian, Y. Rephaeli, A.M. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9327-2
C. Pfrommer, T. Enßlin, Astron. Astrophys. **413**, 17 (2004)
B.M. Poggianti, I. Smail, A. Dressler et al., Astrophys. J. **601**, 197 (2004)
Y. Rephaeli, Astrophys. J. **212**, 608 (1977)
Y. Rephaeli, Astrophys. J. **227**, 364 (1979)
Y. Rephaeli, D. Gruber, Astrophys. J. **595**, 137 (2003)
Y. Rephaeli, D. Gruber, P. Blanco, Astrophys. J. **511**, 21 (1999)
Y. Rephaeli, D. Gruber, Y. Arieli, Astrophys. J. **649**, 673 (2006)
Y. Rephaeli, J. Nevalainen, T. Ohashi, A.M. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9314-7
E. Rizza, C. Loken, M. Bliton, K. Roettiger, J.O. Burns, Astron. J. **119**, 21 (2000)
H.J.A. Röttgering, New Astron. Rev. **47**, 405 (2003)
H.J.A. Röttgering, M.H. Wieringa, R.W. Hunstead, R.D. Ekers, Mon. Not. R. Astron. Soc. **290**, 577 (1997)
L. Rudnick, K.M. Blundell, in *The Riddle of Cooling Flows in Galaxies and Clusters of Galaxies*, ed. by T. Reiprich, J. Kempner, N. Soker (2003a). Available electronically at http://www.astro.virginia.edu/coolflow/proc.php
L. Rudnick, K.M. Blundell, Astrophys. J. **588**, 143 (2003b)
C.L. Sarazin, in *X-ray and Radio Connections*, ed. by L.O. Sjouwerman, K.K. Dyer. Published electronically by NRAO (2005). http://www.aoc.nrao.edu/events/xraydio
C.L. Sarazin, J.C. Kempner, Astrophys. J. **533**, 73 (2000)
J.-L. Sauvageot, E. Belsole, G.W. Pratt, Astron. Astrophys. **444**, 673 (2005)
S. Schindler, in *Merging Processes in Galaxy Clusters*, ed. by L. Feretti, I.M. Gioia, G. Giovannini, Astrophys. Sp. Sc. Lib. (Kluwer) **272**, 229 (2002)
P. Schuecker, H. Böhringer, in *Diffuse Thermal and Relativistic Plasma in Galaxy Clusters*, ed. by H. Böhringer, L. Feretti, P. Schuecker. MPE Report **271**, 43 (1999)
D. Sijbring, A radio continuum and H I line study of the Perseus cluster. PhD Thesis, Groningen, 1993
O.B. Slee, A.L. Roy, M. Murgia, H. Andernach, M. Ehle, Astron. J. **122**, 1172 (2001)

R. Subrahmanyan, A.J. Beasley, W.M. Goss, K. Golap, R.W. Hunstead, Astron. J. **125**, 1095 (2003)
R.A. Sunyaev, Y.B. Zel'dovich, Comments Astrophys. Space Phys. **4**, 173 (1972)
G.B. Taylor, R.A. Perley, Astrophys. J. **416**, 554 (1993)
G.B. Taylor, F. Govoni, S.W. Allen, A.C. Fabian, Mon. Not. R. Astron. Soc. **326**, 2 (2001)
G.B. Taylor, A.C. Fabian, S.W. Allen, Mon. Not. R. Astron. Soc. **334**, 769 (2002)
G.B. Taylor, A.C. Fabian, G. Gentile et al., Mon. Not. R. Astron. Soc. **382**, 67 (2007)
M. Thierbach, U. Klein, R. Wielebinski, Astron. Astrophys. **397**, 53 (2003)
J.P. Vallée, M.J. MacLeod, N.W. Broten, Astron. Astrophys. **156**, 386 (1986)
T. Venturi, S. Giacintucci, G. Brunetti et al., Astron. Astrophys. **463**, 937 (2007)
A. Vikhlinin, M. Markevitch, S.S. Murray, Astrophys. J. **549**, L47 (2001)
C. Vogt, T.A. Enßlin, Astron. Astrophys. **412**, 373 (2003)
C. Vogt, T.A. Enßlin, Astron. Astrophys. **434**, 67 (2005)
H. Völk, A.M. Atoyan, Astrophys. Phys. **11**, 73 (1999)
H. Völk, F.A. Aharonian, D. Breitschwerdt, Space Sci. Rev. **75**, 279 (1996)
M.A.G. Willson, Mon. Not. R. Astron. Soc. **151**, 1 (1970)

Chapter 7
Cosmological Shock Waves

A.M. Bykov · K. Dolag · F. Durret

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9312-9

Abstract Large-scale structure formation, accretion and merging processes, AGN activity produce cosmological gas shocks. The shocks convert a fraction of the energy of gravitationally accelerated flows to internal energy of the gas. Being the main gas-heating agent, cosmological shocks could amplify magnetic fields and accelerate energetic particles via the multi-fluid plasma relaxation processes. We first discuss the basic properties of standard single-fluid shocks. Cosmological plasma shocks are expected to be collisionless. We then review the plasma processes responsible for the microscopic structure of collisionless shocks. A tiny fraction of the particles crossing the shock is injected into the non-thermal energetic component that could get a substantial part of the ram pressure power dissipated at the shock. The energetic particles penetrate deep into the shock upstream producing an extended shock precursor. Scaling relations for postshock ion temperature and entropy as functions of shock velocity in strong collisionless multi-fluid shocks are discussed. We show that the multi-fluid nature of collisionless shocks results in excessive gas compression, energetic particle acceleration, precursor gas heating, magnetic field amplification and non-thermal emission. Multi-fluid shocks provide a reduced gas entropy production and could also modify the observable thermodynamic scaling relations for clusters of galaxies.

Keywords Large scale structure of universe · Shock waves · Acceleration of particles · X-rays: galaxies: clusters

A.M. Bykov (✉)
A.F. Ioffe Institute of Physics and Technology, St. Petersburg 194021, Russia
e-mail: byk@astro.ioffe.ru

K. Dolag
Max-Planck-Institut für Astrophysik, P.O. Box 1317, 85741 Garching, Germany
e-mail: kdolag@mpa-garching.mpg.de

F. Durret
Institut d'Astrophysique de Paris, CNRS, Université Pierre et Marie Curie, 98bis Bd Arago, 75014 Paris, France
e-mail: durret@iap.fr

1 Introduction

The observed large scale structure of the Universe is thought to be due to the gravitational growth of density fluctuations in the post-inflation era. In this model, the evolving cosmic web is governed by non-linear gravitational growth of the initially weak density fluctuations in the dark energy dominated cosmology. The web is traced by a tiny fraction of luminous baryonic matter. Cosmological shock waves are an essential and often the only way to power the luminous matter by converting a fraction of gravitational power to thermal and non-thermal emissions of baryonic/leptonic matter.

At high redshifts ($z > 1100$) the pre-galactic medium was hot, relatively dense, ionised, with a substantial pressure of radiation. The cosmic microwave background (CMB) observations constrain the amplitudes of density inhomogeneities to be very small at the last scattering redshift $z \sim 1000$. Strong non-linear shocks are therefore unlikely at that stage. The universe expands, the matter cools, and eventually recombines, being mostly in neutral phase during the "dark ages" of the universe. At some redshift, $6 < z < 14$, hydrogen in the universe is reionised, likely due to UV radiation from the first luminous objects, leaving the intergalactic medium (IGM) highly reionised (see e.g., Fan et al. (2006) for a recent review). The reionisation indicates the formation of the first luminous objects at the end of the "dark ages", either star-forming galaxies or Active Galactic Nuclei (AGN). The compact luminous objects with an enormous energy release would have launched strong (in some cases, relativistic) shock waves in the local vicinity of the energetic sources. At the same evolution stage, formation of strong density inhomogeneities in the cosmic structure occurs. Since then the non-linear dynamical flows in the vicinity of density inhomogeneities would have created large scale cosmic structure shocks of modest strength, thus heating the baryonic matter and simultaneously producing highly non-equilibrium energetic particle distributions, magnetic fields and electromagnetic emission.

Most of the diffuse X-ray emitting matter was likely heated by cosmological shocks of different scales. Accretion and merging processes produce large-scale gas shocks. Simulations of structure formation in the Universe predict that in the present epoch about 40% of the normal baryonic matter is in the Warm-Hot Intergalactic Medium (WHIM) at overdensities $\delta \sim 5$–10 (e.g., Cen and Ostriker 1999; Davé et al. 2001). The WHIM is likely shock-heated to temperatures of 10^5–10^7 K during the continuous non-linear structure evolution and star-formation processes.

The statistics of cosmological shocks in the large-scale structure of the Universe were simulated in the context of the ΛCDM-cosmology using PM/Eulerian adiabatic hydrodynamic codes (e.g., Miniati et al. 2000; Ryu et al. 2003; Kang et al. 2007) and more recently with a smoothed particle hydrodynamic code by Pfrommer et al. (2006). They identified two main populations of cosmological shocks: (i) high Mach number "external" shocks due to accretion of cold gas on gravitationally attracting nodes, and (ii) moderate Mach number $(2 \leq \mathcal{M}_s \leq 4)$ "internal" shocks. The shocks are due to supersonic flows induced by relaxing dark matter substructures in relatively hot, already shocked, gas. The internal shocks were found by Ryu et al. (2003) to be most important in energy dissipation providing intercluster medium (ICM) heating, and they were suggested by Bykov et al. (2000) to be the likely sources of non-thermal emission in clusters of galaxies.

Hydrodynamical codes deal with N-body CDM and single-fluid gas dynamics. However, if a strong accretion shock is multi-fluid, providing reduced post-shock ion temperature and entropy, then the internal shocks could have systematically higher Mach numbers.

Space plasma shocks are expected to be *collisionless*. Cosmological shocks, being the main gas-heating agent, generate turbulent magnetic fields and accelerate energetic particles

via collisionless multi-fluid plasma relaxation processes thus producing non-thermal components. The presence of these non-thermal components may affect the global dynamics of clusters of galaxies (Ostriker et al. 2005) and the σ_v–T, M–T, L_X–T scaling relations (Bykov 2005). Detailed discussion of the cosmological simulations of the scaling relations with account of only thermal components can be found in Borgani et al. (2008—Chap. 13, this issue).

In Sect. 2 we discuss the basic features of the standard collisional shocks. The main part of the review is devoted to physical properties of cosmological shocks with an accent on collisionless shocks and associated non-thermal components. In Sect. 5 we discuss the most important features of multi-fluid collisionless shocks in the cosmological context including the effects of reduced entropy production, energetic particle acceleration and magnetic field amplification in the shocks.

2 Single-Fluid MHD-Shocks

Shock waves are usually considered as a sharp transition between a macroscopic supersonic (and super-Alfvénic) upstream flow (state 1) and slowed down to a subsonic velocity downstream flow (state 2), providing a mass flow j_n through the shock surface. It is assumed that a gas particle (or an elementary macroscopic fluid cell) is at any instant of time in the *local thermodynamic equilibrium state* corresponding to the instantaneous values of the macroscopic parameters. The Maxwellian distribution of all species is ensured after a few molecular (or Coulomb) collisions have occurred. The macroscopic parameters characterising the state of the gas, such as density, specific internal energy, or temperature, change slowly in comparison with the rates of the relaxation processes leading to thermodynamic equilibrium. We consider here a single-fluid plasma model assuming complete electron–ion relaxation. Under these conditions, in a frame moving with the shock front, with the matter flux across the shock surface $j_n \neq 0$, the conservation laws for mass (in non-relativistic flows), momentum and energy can be written as follows:

$$j_n \left[\frac{\mathbf{B_t}}{\rho}\right] = B_n[\mathbf{u_t}], \tag{1}$$

$$j_n[\mathbf{u_t}] = \frac{B_n}{4\pi j_n}[\mathbf{B_t}], \tag{2}$$

$$\left[\frac{j_n^2}{\rho} + P + \frac{B_t^2}{8\pi}\right] = 0, \tag{3}$$

$$\left[w + \frac{j_n^2}{2\rho^2} + \frac{u_t^2}{2} + \frac{B_t^2}{4\pi\rho} - \frac{B_n}{4\pi j_n}\mathbf{B_t u_t}\right] = 0. \tag{4}$$

Here $\mathbf{U} = (u_n, \mathbf{u_t})$ is the bulk velocity, $w = \varepsilon + P/\rho$ is the gas enthalpy, ε, P, ρ are the internal gas energy, pressure and density respectively. The subscripts n and t are used for the normal and transverse components respectively. We used the standard notations $[A] = A_2 - A_1$ for the jump of a function A between the downstream and upstream regions. In the MHD case the relations equivalent to (1)–(4) were obtained by de Hoffmann and Teller (1950). The equations are valid in the case of a magnetic field frozen-into moving plasma, where $\mathbf{E} = -\frac{\mathbf{U}}{c} \times \mathbf{B}$. A specific feature of MHD shock waves is the so-called coplanarity theorem (e.g., Landau and Lifshitz 1984) saying that the upstream and downstream magnetic

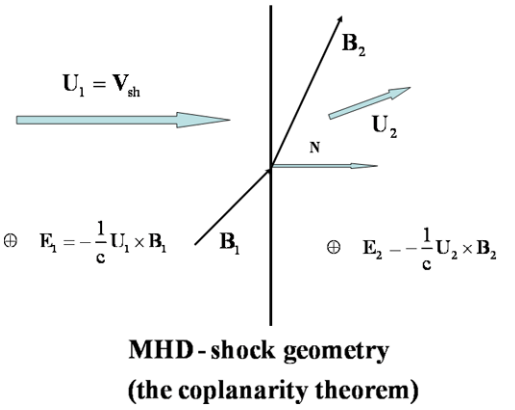


Fig. 1 A sketch illustrating the coplanarity theorem for a plane ideal MHD-shock. The upstream and downstream bulk velocities $\mathbf{U_1}$ and $\mathbf{U_2}$, magnetic fields $\mathbf{B_1}$ and $\mathbf{B_2}$ and the shock normal $\mathbf{N}$ all lie in the same plane. The shock is at rest in the reference frame where also $u_{t1} = 0$. The shock is of infinitesimal width in the sketch. Simulated structure of the transition region of a collisionless shock is shown in Fig. 2 and Fig. 3 where its finite width is apparent

fields $\mathbf{B_1}$ and $\mathbf{B_2}$ and the shock normal all lie in the same plane as it is illustrated in Fig. 1. It is important to note that if $B_n \neq 0$ there is an especial reference frame where local velocity $\mathbf{U}$ and magnetic field $\mathbf{B}$ are parallel both in the upstream and downstream, providing $\mathbf{E} = 0$.

From (1)–(4) one may obtain a generalised Rankine-Hugoniot (RH) adiabat

$$\varepsilon_2 - \varepsilon_1 + \frac{1}{2}\left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right)\{(P_2 + P_1) + \frac{1}{8\pi}(B_{t2} - B_{t1})^2)\} = 0. \tag{5}$$

The RH adiabat connects the macroscopic parameters downstream of the flow once the upstream state is known. In a parallel shock ($\mathbf{B}_t = 0$)

$$r = \frac{\rho_2}{\rho_1} = \frac{(\gamma_g + 1)\mathcal{M}_s^2}{(\gamma_g - 1)\mathcal{M}_s^2 + 2}, \tag{6}$$

$$\frac{T_2}{T_1} = \frac{[2\gamma_g\mathcal{M}_s^2 - (\gamma_g - 1)][(\gamma_g - 1)\mathcal{M}_s^2 + 2]}{(\gamma_g + 1)^2\mathcal{M}_s^2}, \tag{7}$$

where γ_g is the gas adiabatic exponent. We restrict ourselves here to a fast mode shock where $c_{s1} < u_1$, and $v_{a2} < u_2 < c_{s2}$, for $v_{a1} < c_{s1}$. The phase velocity v_{a2} is the Alfvén velocity in the downstream, c_{s1}, c_{s2} are the sound speeds in the upstream and downstream respectively. We define here the shock Mach numbers as $\mathcal{M}_s = v_{sh}/c_{s1}$ and $\mathcal{M}_a = v_{sh}/v_{a1}$.

In the case of a perpendicular shock ($B_n = 0$) the compression ratio is

$$r = \frac{\rho_2}{\rho_1} = \frac{2(\gamma_g + 1)}{\psi + (\psi^2 + 4(\gamma_g + 1)\,(2 - \gamma_g)\,\mathcal{M}_a^{-2})^{1/2}},$$
$$\psi = (\gamma_g - 1) + (2\mathcal{M}_s^{-2} + \gamma_g\,\mathcal{M}_a^{-2}). \tag{8}$$

In a single fluid strong shock with $\mathcal{M}_s \gg 1$ and $\mathcal{M}_a \gg 1$ one gets

$$T_2 \approx 2 \cdot \frac{(\gamma_g - 1)}{(\gamma_g + 1)^2}\,\mu v_{sh}^2 = 1.38 \times 10^7\, v_{s8}^2\ (\mathrm{K}), \tag{9}$$

for any magnetic field inclination (e.g., Draine and McKee 1993). The mass per particle μ was assumed to be $[1.4/2.3]m_H$ and v_{s8} is the shock velocity in 10^8 cm s^{-1}.

The RH adiabat does not depend on the exact nature of the dissipation mechanisms that provide the transition between the states 1 and 2. It assumes a single-fluid motion in regular

electromagnetic fields. However, the dissipative effects control the thickness of the shock transition layer. In the case of a weak shock of Mach number $\mathcal{M}_s - 1 \ll 1$ the thickness is large enough, allowing a macroscopic hydrodynamical description of the fluid inside the shock transition layer (e.g., Landau and Lifshitz 1984). The gas shock width Δ in collisional hydrodynamics without magnetic fields is given by

$$\Delta = \frac{8aV^2}{(P_2 - P_1)\,(\partial^2 V/\partial P^2)_s}. \tag{10}$$

Following Landau and Lifshitz (1959) the gas shock width in (10) can be expressed through the viscosities η and ζ, and thermal conductivity κ, since

$$a = \frac{1}{2\rho v_s^3}\left[\left(\frac{4}{3}\eta + \zeta\right) + \kappa\left(\frac{1}{c_v} - \frac{1}{c_p}\right)\right].$$

Here c_v and c_p are specific heats at constant volume and at constant pressure respectively. Extrapolating (10) to a shock of finite strength where $P_2 - P_1 \sim P_2$, one may show that the gas shock width Δ is of the order of the mean free path λ.

It is instructive to note that the entropy is non-monotonic inside the finite width of a weak gas shock $(\mathcal{M}_s - 1) \ll 1$ and the total RH jump of the entropy Δs across the shock is of the third order in $(\mathcal{M}_s - 1)$:

$$\Delta s = \frac{1}{12T_1}\left(\frac{\partial^2 V}{\partial P^2}\right)_s (P_2 - P_1)^3 \propto (\mathcal{M}_s - 1)^3, \tag{11}$$

while the density, temperature and pressure jumps are $\propto (\mathcal{M}_s - 1)$ (Landau and Lifshitz 1959).

In plasma shocks the shock structure is more complex because of a relatively slow electron–ion temperature relaxation. Such a shock consists of an ion viscous jump and an electron–ion thermal relaxation zone. In the case of plasma shocks the structure of the ion viscous jump is similar to the single fluid shock width structure discussed above and can be studied accounting for the entropy of an isothermal electron fluid. The shock ion viscous jump has a width of the order of the ion mean free path. The scattering length (the mean free path to $\pi/2$ deflection) λ_p of a proton of velocity v_7 (measured in 100 km s^{-1}) due to binary Coulomb collisions with plasma protons of density n (measured in cm^{-3}) can be estimated as $\lambda_p \approx 7 \times 10^{14}\ v_7^4 n^{-1}$ cm (Spitzer 1962). After the reionisation ($z < 6$) the Coulomb mean free path in the WHIM of overdensity δ is $\lambda_p \approx 3.5 \times 10^{21}\ v_7^4 \delta^{-1}(1+z)^{-3}(\Omega_b h^2/0.02)^{-1}$ cm. Here and below Ω_b is the baryon density parameter. The mean free path due to Coulomb collisions is typically some orders of magnitude smaller than that for the charge-exchange collisions in the WHIM after reionisation. The ion-electron thermal relaxation occurs on scales about $\lambda_e \times \sqrt{m_p/Zm_e}$. Since $\lambda_e \sim \lambda_p$, the width of the relaxation zone is substantially larger than the scale size of the ion viscous jump. The application of the single fluid shock model (1)–(4) to electron–ion plasmas assumes full ion–electron temperature relaxation over the shock width. For a discussion of the relaxation processes see e.g., Bykov et al. (2008—Chap. 8, this issue and references therein).

In a rarefied hot cosmic plasma the Coulomb collisions are not sufficient to provide the viscous dissipation of the incoming flow, and collective effects due to the plasma flow instabilities play a major role, providing the *collisionless* shocks, as it is directly observed in the heliosphere. The observed structure of supernova remnants (e.g., Weisskopf and Hughes 2006) is consistent with that expected if their forward shocks are collisionless. Moreover, the

non-thermal synchrotron emission seen in radio and X-rays is rather a strong argument for high energy particle acceleration by the shock that definitely favours its collisionless nature. That allows us to suggest that cosmological shocks in a rarefied highly ionised plasma (after the reionisation epoch) are likely to be collisionless. There are yet very few observational studies of cosmological shocks (e.g., Markevitch and Vikhlinin 2007). We review some basic principles of collisionless shock physics in the next section.

3 Collisionless Shocks

Since the discovery of the solar wind in the early 1960's it has been realised that the rapid rise time of magnetic storms observed in the Earth suggested very thin *collisionless* shocks created by solar flares (see for a discussion Sagdeev 1966; Kennel et al. 1985). The thickness of a viscous jump in a strong *collisional* shock is of the order of a mean free path (see e.g., Zel'dovich and Raizer 1967). The Coulomb collision mean free path in the tenuous solar wind plasma is comparable to the Sun–Earth distance, and thus the magnetic storm rising time due to standard collisional shocks would exceed the observed time by orders of magnitude.

There are very specific features of collisionless plasma shocks (Sagdeev 1966). Shocks in dense enough plasma with frequent Coulomb collisions evolve very fast to Maxwellian particle distributions with very few particles at high energies. On the contrary, in collisionless plasma shocks, a small minority of particles could gain a disproportionate share of the energy and become non-Maxwellian. Collisionless shocks enable acceleration of a small fraction of the particles to very high energies. Moreover, the accelerated particles could carry away a substantial amount of the kinetic energy of the plasma flow dissipated at the shock. The energetic particles can penetrate far into the shock upstream gas, to create an extended shock precursor. The cold gas in the shock upstream is decelerated and pre-heated by the energetic particle and MHD-wave pressure on a scale larger than a mean free path of an energetic particle. This occurs not only at the bow shock of the Earth at moderately low energies, but also in astrophysical shocks at highly relativistic energies (e.g., Russell 2005).

A direct study of collisionless shock waves in a laboratory is an extremely difficult task. Most of the experimental data on collisionless shock physics are coming from space experiments. There are direct observational data on the shock wave structure in the interplanetary medium with clear evidence for ion and electron acceleration by the shocks (e.g., Tsurutani and Lin 1985; Russell 2005).

Computer simulations of the full structure of collisionless shock waves describe the kinetics of multi-species particle flows and magneto-hydrodynamic (MHD) waves in the strongly-coupled system. The problem is multi-scale. It requires a simultaneous treatment of both "microscopic" structure of the subshock at the thermal ion gyroradii scale where the injection process is thought to occur, and an extended "macroscopic" shock precursor due to energetic particles. The precursor scale is typically more than 10^9 times the microscopic scale of the subshock transition region.

Energetic particles could be an essential component in the WHIM and clusters of galaxies. Nonthermal particle acceleration at shocks is expected to be an efficient process at different evolutional stages of clusters. Being the governing process of the supernova remnant collisionless shock formation, nonlinear wave–particle interactions are responsible for both shock heating and compression of the thermal gas, as well as for creation of an energetic particle population.

3.1 Micro Processes in Collisionless Shocks

In the strong enough collisionless shocks (typically with a Mach number above a few) resistivity cannot provide energy dissipation fast enough to create a standard shock transition (e.g., Kennel et al. 1985) on a microscopic scale. Ion instabilities are important in such shocks that are called *supercritical*.

At the microscopic scale the front of a supercritical shock wave is a transition region occupied by magnetic field fluctuations of an amplitude $\delta B/B \sim 1$ and characteristic frequencies of about the ion gyro-frequency. Generation of the fluctuations is due to instabilities in the interpenetrating multi-flow ion movements. The width of the transition region of a quasi-parallel shock wave reaches a few hundred ion inertial lengths defined as $l_i = c/\omega_{\rm pi} \approx 2.3 \times 10^7 n^{-0.5}$ cm. Here $\omega_{\rm pi}$ is the ion plasma frequency and n is the ionised ambient gas number density measured in cm^{-3}. The ion inertial length in the WHIM can be estimated as $l_i \approx 5.1 \times 10^{10} \delta^{-1/2} (1+z)^{-3/2} (\Omega_{\rm b} h^2/0.02)^{-1/2}$ cm, providing the width of the collisionless shock transition region is smaller by many orders of magnitude than the Coulomb mean free path (that is in the kiloparsec range).

The transition region of a quasi-perpendicular shock is somewhat narrower. The wave generation effects at the microscopic scale have been studied in some detail with hybrid code simulations (e.g., Quest 1988). The large-amplitude magnetic field fluctuations in the shock transition region were directly measured in the interplanetary medium (see e.g., Kan et al. 1991).

There are a few ways to simulate numerically the kinetics of the collisionless plasma phenomena. The most comprehensive study of the collisionless shock structure can be performed with the particle-in-cell (PIC) method where all the plasma components are considered as discrete particles in self-consistent fields. The PIC method allows one to resolve electron scale lengths and frequencies, but on the other hand it requires considerable computer resource. A serious constraint on PIC and other plasma particle simulations of collisionless shocks is that they must be done fully in three spatial dimensions (3D). Jones et al. (1998) have proved that PIC simulations with one or more ignorable dimensions artificially confine particles to field lines and particularly eliminate cross-field diffusion. The effect is especially important for simulations of a creation of a superthermal particle population. All three box dimensions must be involved in these simulations. Exact modelling of electron kinetics in collisionless shocks require PIC simulations (e.g., Hoshino and Shimada 2002; Schmitz et al. 2002). On the other hand, the bulk of the energy of a collisionless shock is carried by the ions and velocity relaxation processes are typically longer than the ion gyroperiods. Thus, though the basic shock physics evolve on ion spatial and temporal scales, the electron kinetic description requires fine resolution at electron scales.

A fairly good description of low-frequency processes of the ion dynamics in the shock transition layer can be achieved with hybrid codes (e.g., Winske and Omidi 1996 and references therein). Hybrid code modelling, which interprets protons as particles and electrons as an inertialess liquid, has made it possible to describe some important features of the (sub)shock waves at the microscopic scale of some hundred times the ion inertial length (e.g., Quest 1988; Lembege et al. 2004; Burgess et al. 2005).

A typical initialisation of a shock in the hybrid code simulations is to inject a relatively cold ion beam (say at the right-hand boundary) and to put a particle reflecting wall at the left-hand boundary of a simulation box. In that case the shock is moving, and the available simulation time is limited, given the finite size of the box. The limited simulation time and the particle statistics per cell are challenging the direct modelling of the origin and evolution of the energetic non-thermal particle population in a shock. To increase the statistics the macro-particle splitting method is used (see e.g., Quest 1992; Giacalone and Ellison 2000).

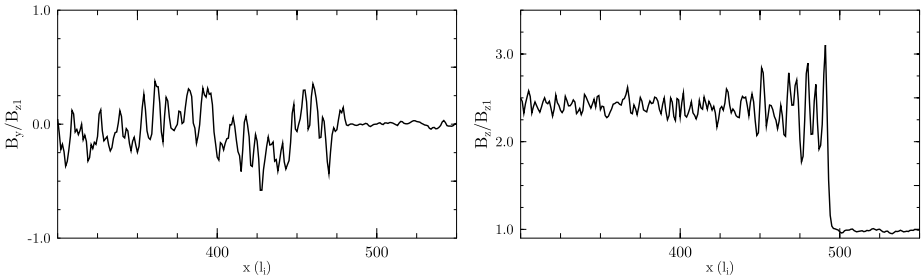


Fig. 2 Hybrid simulated magnetic fields of a quasi-perpendicular shock (80° inclination). The shock propagates along the x-axis, while the initial regular magnetic field is in the x–z plane. We show the B_y and B_z dependence on x in the left and right panels respectively

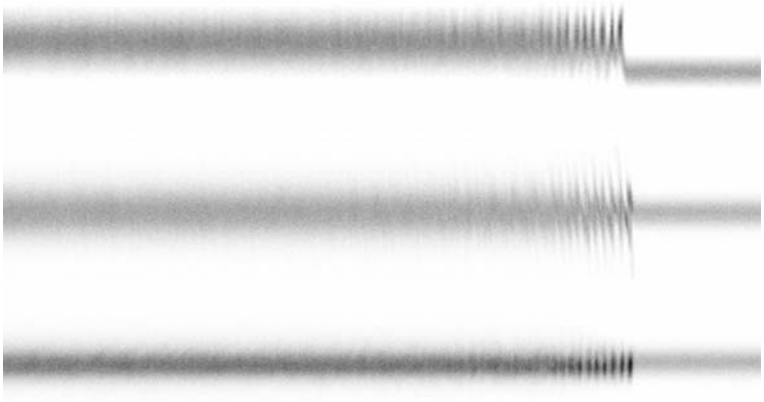


Fig. 3 Hybrid simulated proton phase density in a quasi-perpendicular shock (80° inclination). The shock is moving from left to right in the reference frame where the particle reflecting wall is at rest. The figures show the proton phase densities in v_x–x, v_y–x and v_z–x projections from top to bottom respectively

In Fig. 2 we show the structure of the magnetic field in a quasi-perpendicular shock (inclination angle $\theta_{Bn} \approx 80°$) simulated with a hybrid code for the upstream plasma parameter $\beta \sim 1$. The parameter $\beta = \mathcal{M}_a^2/\mathcal{M}_s^2$, characterises the ratio of the thermal and magnetic pressures. The shock is propagating along the x-axis from the left to the right. The magnetic field is in the x–z plane. The system is periodic in the y-dimension. Phase densities of protons v_x–x, v_y–x, v_z–x are shown in Fig. 3 in the reference frame where the particle reflecting wall (at far left) is at rest while the shock front is moving. The incoming plasma beam in the simulation was composed of protons (90%), alpha particles (9.9%) and a dynamically insignificant fraction of oxygen ions (O VII).

In most of the cases non-relativistic shocks simulated with different hybrid codes had the upstream plasma parameter $\beta \sim 1$. In some cosmological shocks, for example in hot X-ray clusters, plasma parameter β could be $\sim$100 (see Bykov et al. 2008—Chap. 8, this issue). The nature of collisionless shocks in the hot low magnetised plasmas could be different from that in case of $\beta \sim 1$. While the processes of shock formation in the high beta plasmas still require a careful study, there are two experimental studies establishing the existence of the

collisionless shocks for that case. A collisionless shock in a laboratory experiment with a laser-produced ablating plasma of $\beta \sim 300$ was found to have a thickness about 10 c/ω_{pe}, that is orders of magnitude less than the Coulomb mean free path of both ions and electrons in that system (see Bell et al. 1988). In space plasma the terrestrial bow shock under high beta conditions was observed with the *ISEE 1* and *ISEE 2* spacecraft by Farris et al. (1992). These measurements were compared with and found to be in agreement with the predicted values of the Rankine-Hugoniot relations using the simple adiabatic approximation and a ratio of specific heats, gamma, of 5/3. Large magnetic field and density fluctuations were observed, but average downstream plasma conditions well away from the shock were relatively steady, near the predicted Rankine-Hugoniot values. The magnetic disturbances persisted well downstream and a hot, dense ion beam was detected leaking from the downstream region of the shock.

3.2 Heating of Ions in Collisionless Shocks

The heating processes in collisionless shocks are non-trivial. The *irreversible* transformation of a part of the kinetic energy of the ordered bulk motion of the upstream flow into the energy of the random motions of plasma particles in the downstream flow in collisional shocks is due to Coulomb or atomic particle collisions. In collisional non-radiative shocks without slowly relaxing molecular components, the standard single-fluid RH relations are applicable just after a few collisional lengths. The standard single-component shock model predicts a particle temperature $kT = (3/16)mv_{sh}^2$ for $\gamma = 5/3$.

The particle distributions in the collisionless shocks are not Maxwellian. Thus, instead of the standard equilibrium temperature the appropriate moments of the particle distribution function characterising the width of randomised velocity distributions are used. Moreover, the particle velocity distributions are typically anisotropic. It is clearly seen in hybrid simulated proton phase density: in Fig. 3 the velocity distribution widths are different for different projections. One can see in Fig. 3 that a fraction of the incoming ions is reflected by the shock magnetic field jump providing multiple inter-penetrating flows of gyrating ions. Then the field fluctuations randomise the ion phases producing a "coarse-grained" distribution characterised by an effective temperature estimated as the second moment of the velocity distribution.

An analysis of interplanetary collisionless shock observations made with *Advanced Composition Explorer* by Korreck et al. (2007) indicated that quasi-perpendicular shocks are heating ions more efficiently than quasi-parallel shocks. It was also found that effective temperatures of different ions are not necessarily proportional to ion mass, but also depend on the shock inclination angle and plasma parameter β. The widths of the collisionless shocks are extremely narrow (below the astronomical unit) and thus, the observed temperatures would depend on the temperature equilibration processes (both Coulomb and collective) that we will discuss elsewhere (e.g., Bykov et al. 2008—Chap. 8, this issue). The temperature equilibration of different plasma constituents in the WHIM can be studied with spatially resolved spectroscopic observations and thus is a good test for shock models.

3.3 Heating of Electrons in Collisionless Shocks

Electron kinetics in collisionless shocks are different from those of ions. Since most of the observable emission comes from the electrons, they require a careful study. Shocks transfer a fraction of the bulk kinetic energy of the ion flow into large amplitude nonlinear magnetic fluctuations on a short scale of the transition region (see Fig. 2). It is important that

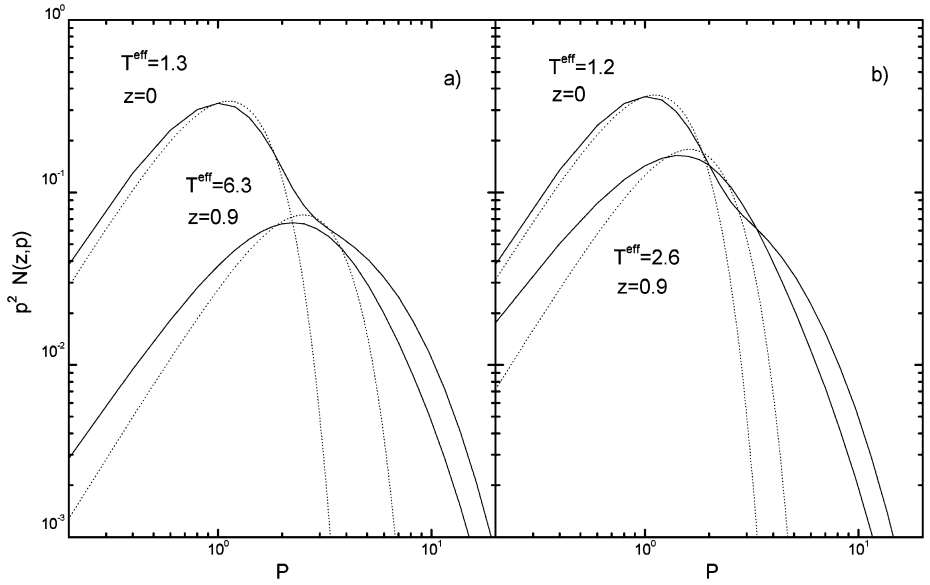


Fig. 4 Electron distribution function simulated in the model of electron heating by strong ion gyroradii scale magnetic fluctuations in a collisionless shock by Bykov and Uvarov (1999). The *left panel* is for $\alpha = \langle(\delta B/B)^2\rangle = 0.4$, the *right panel* for $\alpha = 0.1$

the thermal electron velocities in the ambient medium are higher than the shock speed for a shock Mach number $\mathcal{M}_s < \sqrt{m_p/m_e}$, allowing for a nearly-isotropic angular distribution of the electrons. Non-resonant interactions of these electrons with the large-amplitude turbulent fluctuations in the shock transition region could result in collisionless heating and pre-acceleration of the electrons (Bykov and Uvarov 1999). In Fig. 4 a simulated electron distribution ($p^2N(z, p)$) is shown as a function of the dimensionless electron momentum $p/\sqrt{2m_eT_1}$, where T_1 is the initial electron temperature in the far upstream ($z \to -\infty$). The solid curves are the simulated electron distribution functions at the left boundary ($\tilde{z} = 0$) of the transition region clearly seen in Fig. 2, and at the end of the region ($\tilde{z} = 1$). Dotted lines are the Maxwellian distribution fits allowing to estimate the effective electron temperatures $T^{\rm eff}$ measured relative to T_1. Note that $T^{\rm eff} = 1.2$ at $\tilde{z} = 0$ because of the electron diffusivity effect. One may also clearly see the appearance of non-thermal tails indicating a Fermi type acceleration. It is worth noting that the presence of large-amplitude waves in the shock transition region erodes many of the differences between quasi-parallel and perpendicular shocks, making the electron injection mechanism in that model to be similar for these shocks.

The analysis of observational data on both interplanetary and supernova shocks by Ghavamian et al. (2007) indicates that the electron heating efficiency i.e. T_e/T_i is a declining function of the shock velocity. These authors discussed a model of electron heating with a constant level of electron heating over a wide range of shock speeds (see also Fig. 4 in Bykov and Uvarov 1999), while the ion heating is an increasing function of shock velocity.

3.4 Gas Heating and Entropy Production in Weak Internal Shocks

Heating/acceleration efficiency with weak and moderate strength MHD shocks can be estimated by calculating the energy dissipation rate $\dot{\varepsilon}_h$ of a directed gas motion per unit area of

a weak shock. Defining $\dot{\varepsilon}_{\rm h} = v_{\rm sh}\rho T\Delta s$, where Δs is the difference of the entropies (per unit mass) behind and ahead of the shock front, one can evaluate the thermal gas heating as

$$\dot{\varepsilon}_{\rm h} = (5/4)(\mathcal{M}_{\rm s} - 1)^3 v_{\rm sh}\varepsilon_{\rm T}, \tag{12}$$

where ε_T is the gas internal energy per unit volume (cf. (11)). The energy transfer to the reflected nonthermal particles can be estimated from:

$$\dot{\varepsilon}_{\rm CR} = (\mathcal{M}_{\rm s} - 1)^2 v_{\rm sh}\varepsilon_{\rm B}, \tag{13}$$

where ε_B is the magnetic energy density (Bykov and Toptygin 1987). From (12) and (13) one can see that the gas heating is of third order in $(\mathcal{M}_{\rm s} - 1) \ll 1$ (cf. Landau and Lifshitz 1959), while the wave damping due to the particle acceleration is of second order. Note that in the outer parts of galaxy clusters one would typically expect $\varepsilon_{\rm T} \gg \varepsilon_{\rm B}$. However, the central regions of such a cluster could have $\varepsilon_{\rm T}$ comparable to $\varepsilon_{\rm B}$, as it is the case in the Milky Way. Thus, the weak shocks in the central regions could efficiently accelerate nonthermal particles, reducing the heating of the gas. Particle acceleration by an ensemble of large scale shocks in a cluster of galaxies can create a population of non-thermal particles of sizeable pressure. This may imply a non-steady evolution of non-thermal pressure as modelled by Bykov (2001).

4 Energetic Particle Acceleration in Collisionless Shocks

The reflected ions with a gyro-radius exceeding the width of the shock transition region can then be efficiently accelerated, via the Fermi mechanism, by converging plasma flows carrying magnetic inhomogeneities and MHD waves. In perpendicular shock a net transverse particle momentum gain is due to the work of the electric field on the particle drift motion. The electric field perpendicular to the shock normal exists in all the reference frames for the perpendicular shock. The particle of a momentum **p** crossing back and forth the shock front and being scattered by MHD waves carried with a flow of velocity **u** would undergo a momentum increment $\Delta p \approx \mathbf{p} \cdot \frac{\mathbf{u}}{v} + O((u/v)^2)$ per scattering. A velocity profile in the plane shock is illustrated by the dashed line in Fig. 5 in the test particle case where one neglects the back reaction effect of accelerated particles on the shock. One way to calculate the

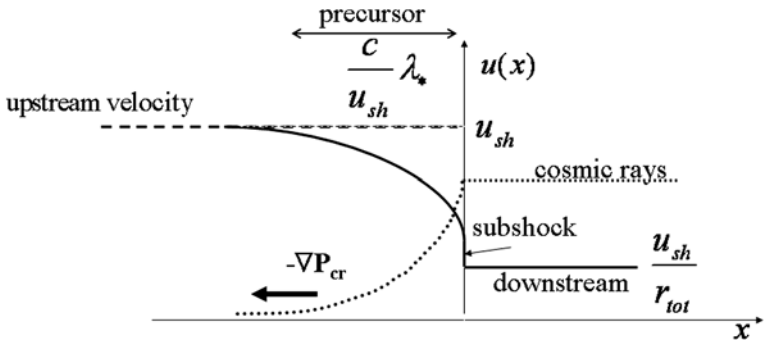


Fig. 5 A sketch illustrating the structure of a cosmic-ray modified shock. The *dashed line* is the shock velocity jump corresponding to the test particle case. The *dotted line* is a spatial distribution of accelerated particles at some momentum $p \gg p_0$. The *solid line* is the CR modified shock velocity profile with the precursor and subshock indicated

accelerated particle spectra in a scattering medium is to use the kinetic equation in the diffusion approximation. The scattering medium in that approach is characterised by momentum dependent particle diffusion coefficients $k_1(p)$ and $k_2(p)$. The shock is considered as a bulk velocity jump (see the dashed line in Fig. 5) assuming that the test particles are injected at $p = p_0$ and the gyroradii of the particles are larger than the shock width. Therefore, in a test particle case particles must be injected at some super-thermal energy to be accelerated by the shock. A solution to the kinetic equation for a nearly isotropic test particle distribution in the phase space is a power-law momentum distribution $f(p, x) \propto (p/p_0)^{-b}$, $p \geq p_0$ where the index

$$b = \frac{3r}{r-1} \tag{14}$$

depends on the shock compression ratio r (Axford et al. 1977; Krymskii 1977; Bell 1978; Blandford and Ostriker 1978). The CR spatial distribution in the model is illustrated in Fig. 5. For a strong shock of $\mathcal{M}_s \gg 1$ and $\mathcal{M}_a \gg 1$ the compression ratio given by (6) and (8) is close to 4 if $\gamma_g = 5/3$ (or even larger if relativistic gas dominates the equation of state). The pressure of the accelerated particles is

$$P_{CR} = \frac{4\pi}{3} \int_{p_0}^{\infty} pvf(p, x) p^2 \, dp. \tag{15}$$

Then for $b = 4$ one may see that $P_{CR} \propto \ln(p_{max}/p_0)$ indicating a potentially large cosmic ray (CR) pressure, if CRs are accelerated to $p_{max} \gg p_0$. The maximal energy of accelerated test particles depends on the diffusion coefficients, bulk velocity and scale-size of the system. The finite scale-size of the shock is usually accounted for by an energy dependent free escape boundary located either in the upstream or in the downstream. For electrons p_{max} can also be limited by synchrotron (or inverse-Compton) losses of relativistic particles.

The test particle shock acceleration time $\tau_a(p)$ can be estimated from the equation

$$\tau_a(p) = \frac{3}{u_{1n} - u_{2n}} \int_{p_0}^{p} \left(\frac{k_1(p)}{u_{1n}} + \frac{k_2(p)}{u_{2n}} \right) \frac{dp}{p} \tag{16}$$

where the normal components of the shock upstream and downstream bulk velocities u_{1n}, u_{2n} are measured in the shock rest frame. Estimations based on a more rigorous approach distinguishing between the mean acceleration time and the variance does not change the results substantially, given the uncertainties in the diffusion model.

5 Cosmic-Ray Modified Multifluid Shocks

The efficiency of the upstream plasma flow energy conversion into nonthermal particles could be high enough providing a hard spectrum of nonthermal particles up to some maximal energy $\varepsilon_\star$. If the efficiency of ram energy transfer to the energetic particles is high enough, an extended shock precursor appears due to the incoming plasma flow deceleration by the fast particle pressure. The precursor scale L is of the order of $(c/v_{sh})\lambda_\star$—orders of magnitude larger than the width of the shock transition region (see Fig. 5). Here $\lambda_\star$ is the maximal mean free path of a particle in the energy-containing part of the spectrum and v_{sh} is the shock velocity. We shall later refer to these energetic particles as cosmic rays.

It has been shown that the front of a strong collisionless shock wave consists of an extended precursor and a viscous velocity discontinuity (subshock) of a local Mach number

that is smaller than the total Mach number of the shock wave (see Fig. 5). The compression of matter at the subshock can be much lower than the total compression of the medium in the shock wave with allowance for high compression in the precursor. We shall refer later to such shocks as CR-modified.

The large scale ("macroscopic") structure of a CR-modified shock can be modelled by a two-fluid approach with a kinetic description of nonthermal particles (see e.g., Blandford and Eichler 1987; Berezhko et al. 1996; Malkov and Drury 2001; Blasi 2004 and references therein) or by a Monte Carlo method (e.g., Jones and Ellison 1991; Ellison et al. 1996). In both methods some suitable parameterisation of particle scattering process must be postulated *a priori*. Monte-Carlo simulations, however, have no assumption of isotropy for particle distributions, and that allows an internally self-consistent treatment of thermal particle injection. While the injection depends on the assumptions made for the particle pitch-angle scattering, these assumptions are applied equally to particles of all energies. The Monte Carlo technique eliminates a free injection parameter, which is present in all models based on the diffusion approximation and is used to set the injection efficiency. The strong feedback between injection, shock structure, and magnetic field amplification makes this property of the Monte Carlo technique particularly important. The Monte Carlo technique allows to iteratively obtain a shock velocity profile and particle distribution function conserving mass, momentum and energy fluxes taking into account the nonlinear feedback from the accelerated energetic particles.

In Fig. 6 Monte Carlo simulated proton spectra (multiplied by $[p/(m_{\rm p}c)]^4$) are shown, in the downstream shock from Vladimirov et al. (2006). To illustrate the dependence of the maximal energy of an accelerated proton on the system scale size, a free escape boundary condition was applied at some distance from the subshock position in the shock rest frame. The heavy solid and dotted curves in the right panel correspond to the free escape boundary located at a distance $10^4 r_{\rm g1}$ (where $r_{\rm g1} = m_{\rm p} v_{\rm sh} {\rm c}/{\rm e}B_1$), the dashed curve has $10^3 r_{\rm g1}$, and the light solid curve has $10^5 r_{\rm g1}$. The simulations were done for a supernova shock in the interstellar medium with a shock speed $v_{\rm sh} = 5000\ {\rm km\ s^{-1}}$ and an unshocked proton number

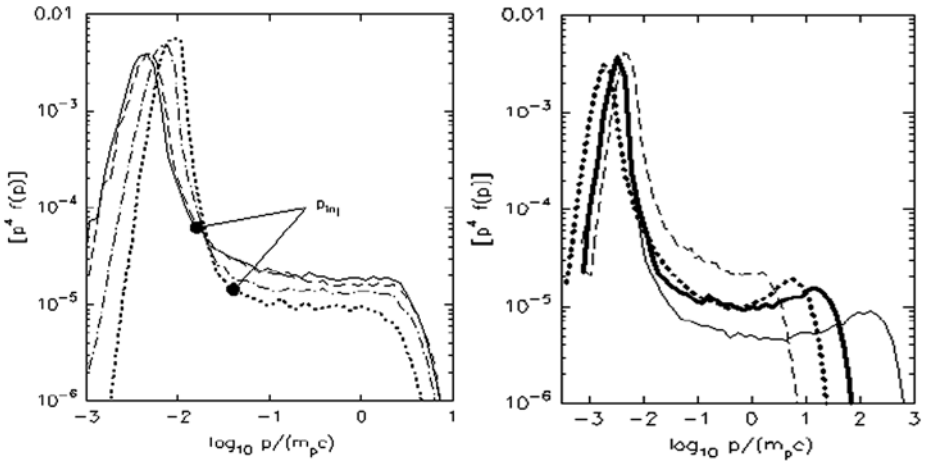


Fig. 6 Spectra of protons accelerated by a strong shock. The spectra were simulated with a non-linear Monte-Carlo model which accounts for particle injection and magnetic field amplification by the shock (for details see Vladimirov et al. 2006). On the *left panel* the transitions from the thermal-like component to the high energy tail are marked by $p_{\rm inj}$. On the *right panel* different curves correspond to different locations of the free escape boundary (see in the text)

density $n_1 = 1\ \mathrm{cm}^{-3}$. In the left panel the spectra are given for the same position of the free escape boundary, but for different prescriptions of the scattering model.

5.1 Magnetic Field Amplification in CR-Dominated Shocks

An important predicted feature of strong shocks with efficient CR acceleration is the possibility to amplify an initial seed magnetic field by orders of magnitude (e.g., Bell and Lucek 2001; Bell 2004). CR current and CR pressure gradient upstream of the strong shock could drive magnetic fluctuations on the shock precursor scale length. The CR-shock precursor scale L is $\sim (c/v_{\mathrm{sh}})\lambda_\star$ which is expected to be above a kpc, moreover, the width is $L \gtrsim 100$ kpc for a shock of a size comparable to that of a galaxy cluster. The precursor scale size L is $\gg 10^9$ times larger than the subshock transition region where strong small scale magnetic field fluctuations are directly produced by instabilities of super-Alfvénic bulk plasma flows illustrated in Fig. 2. That small scale fluctuations are responsible for bulk plasma motion dissipation process and adiabatic amplification of the transverse magnetic field in collisionless shocks. At the same time the collisionless dissipation process is thought to inject a minor fraction of incoming particles to be accelerated to high energies by Fermi mechanism. Recent models of diffusive shock acceleration allows a substantial fraction (say, 30%) of the MHD shock ram pressure to be converted to accelerated particles filling a vicinity of the shock of the scale L. The large scale current and density gradient of the accelerated CRs may convert a fraction of the CR energy to magnetic field due to multi-fluid instabilities of different kinds providing a way to amplify the initial magnetic field by a factor larger than the shock compression ratio.

Recent non-linear simulations of magnetic field amplification in diffusive shock acceleration by a Monte-Carlo model (Vladimirov et al. 2006) and a kinetic model (Amato and Blasi 2006) confirmed the possibility of a significant effect. The amplitude of the fluctuating magnetic field energy density $\mathrm{W_B}$ is of the order of the shock accelerated CR pressure which is in turn a substantial fraction of the shock ram pressure $0.5\rho_1 v_{\mathrm{sh}}^2$. Here ρ_1 is the shock upstream ambient gas density.

For typical cluster parameters the discussed mechanism could provide a μG range magnetic field amplitude in a hundred kpc range scale of CR-modified shock precursor. The Faraday rotation measure RM provided by a strong CR-dominated shock in a cluster can reach values of $\gtrsim 10\ \mathrm{rad\,m}^{-2}$ and even a few times higher. For the case of the so-called Bohm diffusion model the rotation measure RM is proportional to the maximal energy of the ions in the energy-containing part of the CR-spectrum accelerated by the shock. Radio observations, Faraday rotation and synchrotron-Compton emission measurements are used to estimate the magnetic fields in clusters (e.g., Carilli and Taylor 2002; Newman et al. 2002). Large filaments of polarised radio emission of scale size about 400 kpc were discovered by Govoni et al. (2005) in the halo of the cluster of galaxies Abell 2255 and by Bagchi et al. (2006) in Abell 3376 (see Fig. 8). They could be connected to large scale shocks due to accretion/merging activity of the cluster.

5.2 Gas Heating and Entropy Production in Strong CR-Modified Shocks

An exact modelling of a collisionless shock structure taking into account the nonthermal particle acceleration effect requires the nonperturbative self-consistent description of *a multi-component and multi-scale system* including strong MHD-turbulence dynamics. Such a modelling is not feasible at the moment. Instead, a simplified description of a multi-fluid strong shock structure can be used with an appropriate parameterisation of the extended

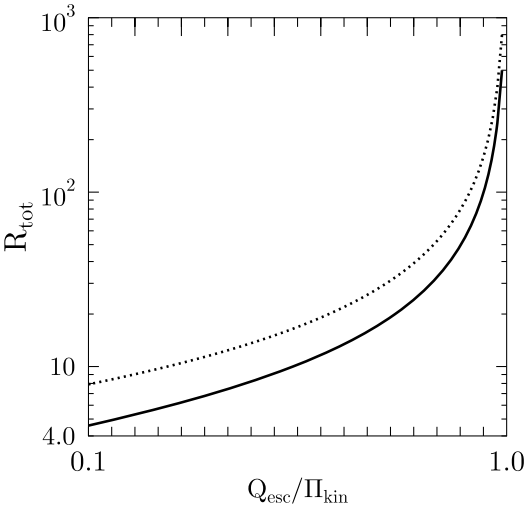


Fig. 7 Total compression ratio $r_{\rm tot}$ of a strong MHD shock modified by efficient particle acceleration as a function of the energy escape flux $Q_{\rm esc}/\Pi_{\rm kin}$ carried by energetic particles, where $\Pi_{\rm kin} = \rho_1 v_{\rm sh}^3/2$. The *upper curve* (*dotted*) corresponds to an effective adiabatic exponent $\gamma = 4/3$ (relativistic gas), while the *lower* (*solid*) *curve* corresponds to $\gamma = 5/3$

pre-shock and of the gas subshock. The predicted observable characteristics of the shocks can be confronted to the observational data. We will now discuss the effects of plasma heating by modified shocks and then make some specific predictions for possible observational tests.

In the shocks with efficient high energy particle acceleration the energy flux carried away by escaping energetic particles $Q_{\rm esc}$ must be accounted for in the energy continuity equations. The energy loss results in a lower effective adiabatic index, but it allows to increase the total compression of the gas in the shock downstream.

The total compression ratio $r_{\rm tot}$ of a strong MHD shock modified by an efficient nonthermal particle acceleration can be estimated as

$$r_{\rm tot} = \frac{\gamma + 1}{\gamma - \sqrt{1 + 2(\gamma^2 - 1) Q_{\rm esc}/\rho_a v_{\rm sh}^3}}, \tag{17}$$

assuming that the energy density in the shock upstream is dominated by the ram pressure and that the CR escape is through the cut-off momentum regime (e.g., Malkov and Drury 2001). Here γ is the effective adiabatic exponent. In Fig. 7 we illustrate the dependence of the compression ratio on $Q_{\rm esc}/\rho_a v_{\rm sh}^3$ for $\gamma = 4/3$ and $5/3$ assuming that the effective adiabatic exponent is between the two values depending on the spectrum of the accelerated relativistic particles.

The distribution function of nonthermal particles and the bulk flow profile in the shock upstream region are sensitive to the total compression ratio $r_{\rm tot}$. Thus, the exact calculation of the escape flux $Q_{\rm esc}$ can be performed only in fully nonlinear kinetic simulations. Nevertheless, an approximate iterative approach (e.g., in the Monte Carlo model discussed above) can be used to make the steady-state distribution function consistent with the shock compression *assuming some diffusion model*. The subshock is the standard gas viscous shock of a Mach number $\mathcal{M}_{\rm sub}$. For that simplified *two-fluid* model of a strong CR-modified shock the effective ion temperature in the downstream $T_{\rm i}^{(2)}$ can be estimated for a shock of a given velocity, if $r_{\rm tot}$ and $r_{\rm sub}$ are known:

$$T_{\rm i}^{(2)} \approx \phi(\mathcal{M}_{\rm sub}) \frac{\mu v_{\rm sh}^2}{\gamma_{\rm g} r_{\rm tot}^2(v_{\rm sh})}, \quad \text{where} \quad \phi(\mathcal{M}_{\rm sub}) = \frac{2\gamma_{\rm g}\mathcal{M}_{\rm sub}^2 - (\gamma_{\rm g} - 1)}{(\gamma_{\rm g} - 1)\mathcal{M}_{\rm sub}^2 + 2}. \tag{18}$$

Single fluid strong shock heating represents the limit $\mathcal{M}_{\rm sub} = \mathcal{M}_{\rm s} \gg 1$, since there is no precursor in that case, resulting in (9). In single-fluid systems the compression ratio $r_{\rm tot} = r_{\rm sub} \to (\gamma_{\rm g}+1)/(\gamma_{\rm g}-1)$ does not depend on the shock velocity and (18) reduces to (9). However, in multi-fluid shocks the total compression ratio depends on the shock velocity and could be substantially higher than that in the single-fluid case. This implies somewhat lower postshock ion temperatures for the strong multi-fluid shock of the same velocity and could be tested observationally. It is convenient to introduce the scaling $r_{\rm tot}(v_{\rm sh}) \propto v_{\rm sh}^{\xi}$ to describe the different cases of strong shock heating (see Bykov 2005 for details). Then from Eq. 18, $T_{\rm i2} \propto \phi(\mathcal{M}_{\rm sub}) v_{\rm sh}^{2(1-\xi)}$. The subshock Mach number $\mathcal{M}_{\rm sub}$ depends, in general, on $\mathcal{M}_{\rm s}$ and $\mathcal{M}_{\rm a}$. Thus, an index σ approximates the velocity dependence of $\phi(\mathcal{M}_{\rm sub}) \propto v_{\rm sh}^{\sigma}$. Finally, if $T_{\rm i2} \propto v_{\rm sh}^{a}$, then the index $a = 2(1-\xi)+\sigma$. For the case of shock precursor heating by CR generated Alfvén waves, the index $a \approx 1.25$ (Bykov 2005).

A distinctive feature of multi-fluid shocks is their high gas compression $r_{\rm tot}(v_{\rm sh})$ that could be well above the single fluid shock limit $(\gamma_{\rm g}+1)/(\gamma_{\rm g}-1)$ (see Fig. 7). At the same time entropy production for a strong multi-fluid shock scales as $[r_{\rm tot}(v_{\rm sh})]^{-(\gamma_{\rm g}+1)}$ and it is significantly reduced compared to the single-fluid shock of the same velocity. The effects are due to energetic particle acceleration and magnetic field amplification.

Energetic particles penetrate into the shock upstream region. They are coupled with the upstream gas through fluctuating magnetic fields (including the Alfvén waves generated by the energetic particles). Magnetic field dissipation provides gas preheating and entropy production in the extended shock precursor. Such a heated pre-shock region of $kT \lesssim 0.5$ keV would appear as an extended filament of width $L \sim (c/v_{\rm sh})\lambda_{\star} \gtrsim 3 \times 10^{14} \epsilon_{\star} B_{-6}^{-1}$ cm. Here $\epsilon_{\star}$ (in GeV) is the highest energy of the hard branch of the accelerated particle spectrum. If $B_{-6} \sim 0.1$ in the cluster outskirts and if the hard spectrum of energetic nuclei extends to $\sim 10^9$ GeV (cf. Norman et al. 1995) we have $L \sim 1$ Mpc and even wider. Projected on a hot X-ray cluster, such filaments could produce a soft X-ray component "excessive" to that produced by the hot cluster. A warm gas (~ 0.2 keV) emission filament found with *XMM-Network* in the outskirts of the Coma cluster by Finoguenov et al. (2003) could be an extended heated precursor of a strong multi-fluid accretion shock. For a detailed review of the soft X-ray/EUV excesses see Durret et al. (2008—Chap. 4, this issue).

5.3 ICM Entropy Production by Multifluid Accretion Shocks

Cold gas falling into the dark matter (DM) dominated gravitational well passes through a strong accretion shock. The shock is a source of gas entropy production in the intercluster medium (ICM) (e.g., Knight and Ponman 1997; Tozzi and Norman 2001; Voit et al. 2003). The post-shock entropy $K = K_{\rm b} T/\rho^{2/3}$ used in the ICM analysis and simulations (e.g., Bialek et al. 2001) is related to the standard thermodynamic entropy s through $K \propto \exp(s/c_{\rm v})$. In the standard scenario with a single-fluid accretion shock the post-shock entropy scales $K_{\rm sf} \propto v_{\rm sh}^2 \rho_1^{-2/3}$ (e.g., Voit et al. 2003).

The multi-fluid nature of the collisionless accretion shock modifies the standard scaling relation to be

$$K_{\rm mf} \propto v_{\rm sh}^2 [r_{\rm tot}(v_{\rm sh})]^{-(1+\gamma_{\rm g})} \phi(\mathcal{M}_{\rm sub}) \rho_1^{(1-\gamma_{\rm g})}. \tag{19}$$

The compression ratio in CR-shocks is higher than in a strong single-fluid shock of the same velocity resulting in reduced post-shock entropy production. For example, in the case of Alfvén heating the post-shock entropy of a multi-fluid shock reduces as $K_{\rm mf}/K_{\rm sf} \sim (15/\mathcal{M}_{\rm a})$ for $\mathcal{M}_{\rm a} > 15$ and $\mathcal{M}_{\rm s}^2 > \mathcal{M}_{\rm a}$. Here and below in numerical estimations we assume $\gamma_{\rm g} = 5/3$, though a non-thermal baryonic component could reduce the index $\gamma_{\rm g}$.

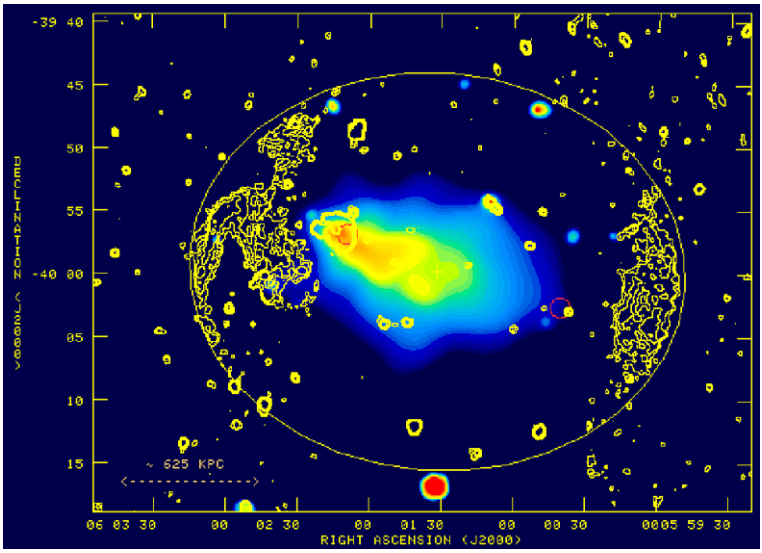


Fig. 8 Abell 3376 (taken from Bagchi et al. 2006: X-ray emission from XMM-Newton archive data, with VLA 1.4 GHz radio contours superimposed. The *ellipse* shows an elliptical fit to the peripheral radio structures, and the '+' marks its centre. The *circles* mark the positions of the two brightest cluster galaxies

Since $r_{\rm tot}(v_{\rm sh})$ and $\phi(\mathcal{M}_{\rm sub})$ are shock velocity dependent, the simple scaling $K \propto v_{\rm sh}^2 \rho_1^{-2/3}$ is not valid. In CR-modified shocks $K_{\rm mf} \propto v_{\rm sh}^{\nu} \rho_1^{(1-\gamma_{\rm g})}$ or $K_{\rm mf} \propto T^{\nu/a}$, where $\nu = 2 - (1+\gamma_{\rm g})\xi + \sigma$. For the case of Alfvén wave heating the index ν is $\lesssim 1.25$ and $K_{\rm mf}$ is $\propto T^{0.8}$ assuming $\gamma_{\rm g} = 5/3$. Recently Ponman et al. (2003) and Piffaretti et al. (2005) found that the dispersion in the observed cluster entropy profiles is smaller if an empirical relation $K \propto T^{0.65}$ is used instead of the standard $K \propto T$ (see also Pratt et al. 2006).

Consider the simple model of smooth accretion of cold gas through a strong accretion shock by Voit et al. (2003). The gas of velocity $v_{\rm ac}$ accretes at a rate $\dot{M}_g$ through the shock at a radius $r_{\rm ac}$ where

$$\dot{M}_g = 4\pi r_{\rm ac}^2 \rho_1 v_{\rm ac}, \quad v_{\rm ac}^2 = 2GM\xi r_{\rm ac}^{-1}, \quad \xi = 1 - r_{\rm ac}/r_{\rm ta}. \tag{20}$$

Here $M(t)$ is the cluster mass and $r_{\rm ta}$ is the matter turnaround radius. Then the entropy $K_{\rm mf}$ just behind the multi-fluid shock is expressed through $T_i^{(2)}(v_{\rm ac})$ and $\rho_2 = r_{\rm tot}(v_{\rm ac})\rho_1$. In the Alfvén wave heating case $K_{\rm mf}(t) \propto (Mt)^{(1+\sigma)/3}$, instead of $K_{\rm sf}(t) \propto (Mt)^{2/3}$ in the single-fluid regime. A multi-fluid shock results in a slower post-shock entropy production. As we have noted above, the regime of CR-shock compression depends on the plasma parameter β in the infalling gas. The plasma parameter β is currently poorly known because the intercluster magnetic fields are not well constrained. The effects of shock modifications are important for both the models of smooth accretion of cold gas and for accretion of hierarchical structures.

Preheating of accreting gas by different physical processes (e.g., due to early star formation in a protocluster region) was suggested by Evrard and Henry (1991), as a possible reason for the breaking of the scaling relations for pure gravitational cluster compression by Kaiser (1986). The observed high metallicity of clusters at different redshifts indicates that strong starburst activity was highly likely at some stage. The preheating produces some initial level of gas entropy ("entropy floor", see e.g., extensive simulations by Bialek et al. 2001;

Borgani et al. 2001; Borgani et al. 2005). Multi-fluid strong shocks provide a natural alternative way of preheating accreting gas. The non-thermal components are essential for detailed modelling of global properties of X-ray clusters, including the mass-temperature and luminosity-temperature relations (Ostriker et al. 2005).

6 Shocks in Large-Scale Structure

Simulations of the cosmic large-scale structure (LSS) predict that about 40–50% of baryons at epoch $z < 2$ could reside in the Warm-Hot Intergalactic Medium (WHIM) with temperatures 10^5–10^7 K at moderate overdensities $\delta \sim 10$ (Cen and Ostriker 1999; Davé et al. 2001). The WHIM heating is due to shocks driven by gravitationally accelerated flows in the LSS structure formation scenario (e.g., Kang et al. 2007). The simulations demonstrate that the X-ray and ultraviolet O VI, O VII and O VIII lines and the H I Lyα line traces the low-density cosmic web filamentary structures. Intervening metal absorption systems of highly ionised C, N, O, Ne in the soft X-ray spectra of bright AGN were suggested to trace the WHIM. The detection of shocked WHIM requires very sensitive UV and X-ray detectors, both for absorption and for emission processes (see, e.g., Kaastra et al. 2008; Richter et al. 2008—Chaps. 9 and 3, this issue). Dedicated future missions like the *Cosmic Origin Spectrograph* (*COS*), the X-Ray Evolving Universe Spectrometer (*XEUS*), *Constellation-X* and the Diffuse Intergalactic Oxygen Surveyor (*DIOS*) will provide high resolution spectroscopy of the shocked WHIM. The WHIM ions of different charge states have highly non-equilibrium (anisotropic) initial states just behind a collisionless shock that relaxes to equilibrium states through Coulomb collisions. As was discussed above a strong collisionless shock could generate a spectrum of MHD-fluctuations. These MHD-fluctuations can carry a substantial fraction of the shock ram pressure. The velocity fluctuations will result in non-thermal broadening of the lines, potentially important for simulations of emission/absorption spectra of the WHIM and observational data analysis. Specific features of collisionless shock heating of the WHIM ions are discussed in Bykov et al. (2008—Chap. 8, this issue). In this paper we discuss only a few observations of clusters of galaxies.

6.1 Evidence for Shocks in Galaxy Clusters

Clusters of galaxies are believed to form within the hierarchical build up of the large scale structure of the Universe. Small objects collapse first and then merge in a complex manner to form larger and larger structures. Therefore, once in a while during their formation, clusters of galaxies undergo so called major merger events. In such events, proto-cluster structures of similar masses (typical ratios 1:10–1:1) are colliding with super sonic velocities (typically several 1000 km s^{-1}). These merging events are a source of shocks and turbulence. They redistribute and amplify magnetic fields, and they are a source of acceleration of relativistic particles within the intracluster medium. Additionally, it is expected that the accretion of the diffuse, unprocessed (and therefore relatively cold) matter onto the DM node of the cosmic web creates a virialisation shock (also called accretion shock), which is expected to be located far in the cluster periphery (typically a few Mpc from the cluster centre for massive systems). Some examples of relevant cosmological simulations are presented by Dolag et al. (2008—Chap. 15, this issue).

X-ray observations, revealing the thermodynamical state of the intracluster medium are therefore the natural means for searching for the signatures of such non thermal phenomena. However, due to biases in the observational processes, caused by the complex temperature

structure of the intracluster medium, such signatures are very hard to detect (for a more detailed discussion see Mazzotta et al. (2004) and references therein). Nevertheless, some detections of shocks in galaxy clusters have been revealed by high resolution *Chandra* and *XMM-Newton* observations. For a recent review see Markevitch and Vikhlinin (2007). One of the most spectacular examples of a merging galaxy clusters is the case of 1E 0657−56 (Markevitch et al. 2002).

The sensitivities of current X-ray instruments are not sufficient to map the state and structure of the intracluster medium in the periphery of galaxy clusters. However, the discovery of arc like radio emission in the periphery of some clusters (so called radio relics), are thought to trace shocks running through the intracluster medium. Spectacular examples are Abell 3667 (see also Ferrari et al. 2008—Chap. 6, this issue) or the radio relics in Abell 3376 recently discovered by Bagchi et al. (2006), see Fig. 8. Thereby, such radio observations are currently the only possibility to observe shocks outside the central regions of galaxy clusters. A more detailed discussion of numerical models of such radio relics is provided by Dolag et al. (2008—Chap. 15, this issue).

Shocks are also expected to trigger star formation, as indicated by numerical simulations (e.g., Bekki 1999), which will leave detectable imprints in the intracluster medium even long (several Gyr) after the shock passed through. Multiple supernova explosions in the star forming regions (superbubbles) will additionally produce copious small scale shocks and accelerate non-thermal particles (e.g., Bykov 2001). One of the tracers for these processes can be excess metallicity in the intracluster medium, produced by the enhanced star formation period (e.g., Schindler et al. 2005). Figure 9 shows the inferred metallicity map for Abell 3376, indicating previous merger activity of the cluster (see Bagchi et al. 2006).

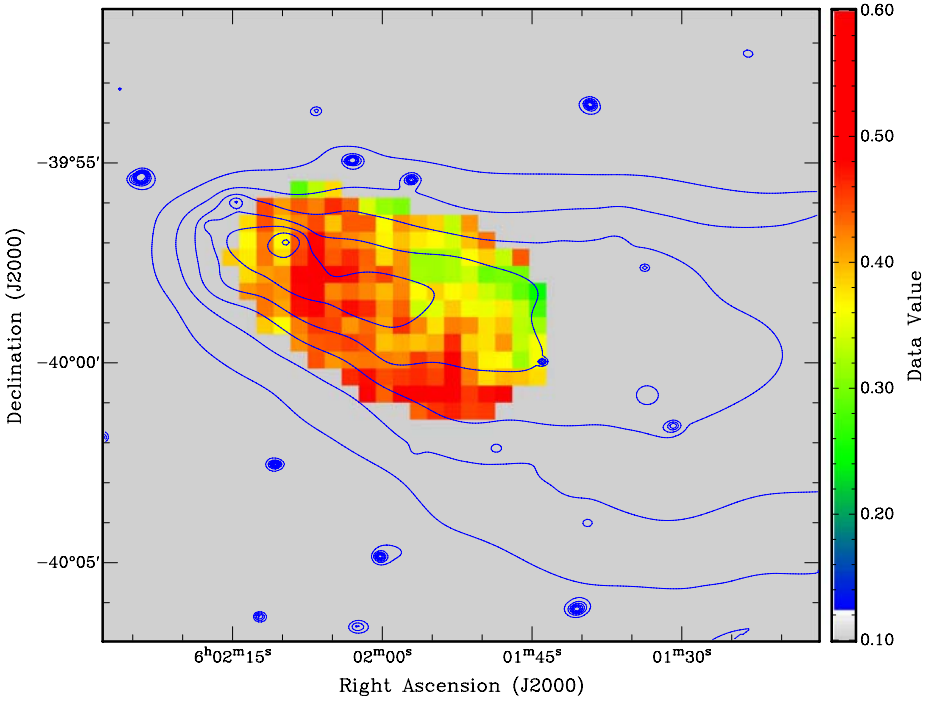


Fig. 9 Metallicity map of the intracluster gas in Abell 3376 (taken from Bagchi et al. 2006)

7 Summary

Cosmological shocks convert a fraction of the energy of gravitationally accelerated flows to internal energy of the gas. They heat and compress the gas and can also accelerate energetic non-thermal particles and amplify magnetic fields. We discussed some specific features of cosmological shocks.

• The standard Rankine-Hugoniot relations based on the conservation laws for a steady single-fluid MHD shock allow to calculate the state of the fluid behind the shock once the upstream state and the shock strength are known. The coplanarity theorem for a plane ideal MHD shock states that the upstream and downstream bulk velocities, magnetic fields and the shock normal all lie in the same plane.

• Cosmological plasma shocks are likely to be collisionless as many other astrophysical shocks observed in the heliosphere and in supernova remnants. We review the basic plasma processes responsible for the microscopic structure of collisionless shocks.

• Collisionless shock heating of ions results in a non-equilibrium state just behind a very thin magnetic ramp region with a strongly anisotropic quasi-Maxwellian ion distributions. The possibility of collisionless heating of electrons by electromagnetic fluctuations in the magnetic ramp region depends on the extension of the fluctuation spectra to the electron gyro-scales, and could depend on the shock Mach number. Then the Coulomb equilibration processes are operating on the scales much larger than the collisionless shock width.

• Extended MHD shock waves propagating in turbulent media could accelerate energetic particles both by Fermi type acceleration in converging plasma flows and by DC electric field in quasi-perpendicular shocks. If the acceleration is efficient, then the strong shock could convert a substantial fraction (more than 10%) of the power dissipated by the upstream bulk flow to energetic particles (cosmic rays). The compression ratio $r_{\rm tot}$ at such a shock can be much higher, while the ion temperature behind the shock $\propto r_{\rm tot}^{-2}$ and the post-shock entropy are lower, than that in a standard single fluid shock. The shock structure consists of an extended precursor and a viscous velocity jump (subshock) indicated in Fig. 5.

• Strong collisionless plasma shocks with an efficient Fermi acceleration of energetic particles could generate strong MHD waves in the upstream and downstream regions and strongly amplify the upstream magnetic fields. A distinctive feature of the shock is a predicted possibility of gas pre-heating in the far upstream region due to MHD wave dissipation, that can produce an extended filament of temperature $\gtrsim 0.1$ keV.

• Shock waves both from the cosmic web formation processes and those due to cluster merging activity can play an important role in clusters of galaxies. Direct evidences for such shocks, as traced by radio relics and the temperature jumps in X-ray observations have been found only in a small number of clusters, and thus we need more observations.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialised X-ray components in clusters of galaxies". A.M.B. thanks M.Yu. Gustov for his help with hybrid shock simulations. He acknowledges the RBRF grant 06-02-16844, a support from RAS Presidium Programmes. A support from NASA ATP (NNX07AG79G) is acknowledged.

References

E. Amato, P. Blasi, Mon. Not. R. Astron. Soc. **371**, 1251 (2006)
W.I. Axford, E. Leer, G. Skadron, in *International Cosmic Ray Conference*, vol. 11 (1977), p. 132
J. Bagchi, F. Durret, G.B.L. Neto, S. Paul, Science **314**, 791 (2006)
K. Bekki, Astrophys. J. **510**, L15 (1999)
A.R. Bell, Mon. Not. R. Astron. Soc. **182**, 147 (1978)

A.R. Bell, Mon. Not. R. Astron. Soc. **353**, 550 (2004)
A.R. Bell, P. Choi, A.E. Dangor, O. Willi, D.A. Bassett, Phys. Rev. A **38**, 1363 (1988)
A.R. Bell, S.G. Lucek, Mon. Not. R. Astron. Soc. **321**, 433 (2001)
E.G. Berezhko, V.K. Elshin, L.T. Ksenofontov, J. Exp. Theor. Phys. **82**, 1 (1996)
J.J. Bialek, A.E. Evrard, J.J. Mohr, Astrophys. J. **555**, 597 (2001)
R. Blandford, D. Eichler, Phys. Rep. **154**, 1 (1987)
R.D. Blandford, J.P. Ostriker, Astrophys. J. **221**, L29 (1978)
P. Blasi, Astropart. Phys. **21**, 45 (2004)
S. Borgani, A. Diaferio, K. Dolag, S. Schindler, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9317-4
S. Borgani, A. Finoguenov, S.T. Kay et al., Mon. Not. R. Astron. Soc. **361**, 233 (2005)
S. Borgani, F. Governato, J. Wadsley et al., Astrophys. J. **559**, L71 (2001)
D. Burgess, E.A. Lucek, M. Scholer et al., Space Sci. Rev. **118**, 205 (2005)
A. Bykov, F.B.S. Paerels, V. Petrosian, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9309-4
A.M. Bykov, Space Sci. Rev. **99**, 317 (2001)
A.M. Bykov, Adv. Space Res. **36**, 738 (2005)
A.M. Bykov, H. Bloemen, Y.A. Uvarov, Astron. Astrophys. **362**, 886 (2000)
A.M. Bykov, I.N. Toptygin, Astrophys. Space Sci. **138**, 341 (1987)
A.M. Bykov, Y.A. Uvarov, J. Exp. Theor. Phys. **88**, 465 (1999)
C.L. Carilli, G.B. Taylor, Annu. Rev. Astron. Astrophys. **40**, 319 (2002)
R. Cen, J.P. Ostriker, Astrophys. J. **514**, 1 (1999)
R. Davé, R. Cen, J.P. Ostriker et al., Astrophys. J. **552**, 473 (2001)
F. de Hoffmann, E. Teller, Phys. Rev. **80**, 692 (1950)
K. Dolag, S. Borgani, S. Schindler, A. Diaferio, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9316-5
B.T. Draine, C.F. McKee, Annu. Rev. Astron. Astrophys. **31**, 373 (1993)
F. Durret, J.S. Kaastra, J. Nevalainen, T. Ohashi, N. Werner, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9313-8
D.C. Ellison, M.G. Baring, F.C. Jones, Astrophys. J. **473**, 1029 (1996)
A.E. Evrard, J.P. Henry, Astrophys. J. **383**, 95 (1991)
X. Fan, C.L. Carilli, B. Keating, Annu. Rev. Astron. Astrophys. **44**, 415 (2006)
M.H. Farris, C.T. Russell, M.F. Thomsen, J.T. Gosling, J. Geophys. Res. **97**, 19121 (1992)
C. Ferrari, F. Govoni, S. Schindler, A. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9311-x
A. Finoguenov, U.G. Briel, J.P. Henry, Astron. Astrophys. **410**, 777 (2003)
P. Ghavamian, J.M. Laming, C.E. Rakowski, Astrophys. J. **654**, L69 (2007)
J. Giacalone, D.C. Ellison, J. Geophys. Res. **105**, 12541 (2000)
F. Govoni, M. Murgia, L. Feretti et al., Astron. Astrophys. **430**, L5 (2005)
M. Hoshino, N. Shimada, Astrophys. J. **572**, 880 (2002)
F.C. Jones, D.C. Ellison, Space Sci. Rev. **58**, 259 (1991)
F.C. Jones, J.R. Jokipii, M.G. Baring, Astrophys. J. **509**, 238 (1998)
J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-y
N. Kaiser, Mon. Not. R. Astron. Soc. **222**, 323 (1986)
J.R. Kan, L.H. Lyu, M.E. Mandt, Space Sci. Rev. **57**, 201 (1991)
H. Kang, D. Ryu, R. Cen, J.P. Ostriker, Astrophys. J. **669**, 729 (2007)
C.F. Kennel, J.P. Edmiston, T. Hada, Geophys. Monogr. Ser. **34**, 1 (1985)
P.A. Knight, T.J. Ponman, Mon. Not. R. Astron. Soc. **289**, 955 (1997)
K.E. Korreck, T.H. Zurbuchen, S.T. Lepri, J.M. Raines, Astrophys. J. **659**, 773 (2007)
G.F. Krymskii, Soviet Physics Doklady **22**, 327 (1977)
L.D. Landau, E.M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1959)
L.D. Landau, E.M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd edn. (Pergamon, London, 1984)
B. Lembege, J. Giacalone, M. Scholer et al., Space Sci. Rev. **110**, 161 (2004)
M.A. Malkov, L.O.C. Drury, Reports Prog. Phys. **64**, 429 (2001)
M. Markevitch, A.H. Gonzalez, L. David et al., Astrophys. J. Lett. **567**, L27 (2002)
M. Markevitch, A. Vikhlinin, Phys. Rep. **443**, 1 (2007)
P. Mazzotta, E. Rasia, L. Moscardini, G. Tormen, Mon. Not. R. Astron. Soc. **354**, 10 (2004)
F. Miniati, D. Ryu, H. Kang et al., Astrophys. J. **542**, 608 (2000)
W.I. Newman, A.L. Newman, Y. Rephaeli, Astrophys. J. **575**, 755 (2002)
C.A. Norman, D.B. Melrose, A. Achterberg, Astrophys. J. **454**, 60 (1995)
J.P. Ostriker, P. Bode, A. Babul, Astrophys. J. **634**, 964 (2005)
C. Pfrommer, V. Springel, T.A. Enßlin, M. Jubelgas, Mon. Not. R. Astron. Soc. **367**, 113 (2006)

R. Piffaretti, P. Jetzer, J.S. Kaastra, T. Tamura, Astron. Astrophys. **433**, 101 (2005)
T.J. Ponman, A.J.R. Sanderson, A. Finoguenov, Mon. Not. R. Astron. Soc. **343**, 331 (2003)
G.W. Pratt, M. Arnaud, E. Pointecouteau, Astron. Astrophys. **446**, 429 (2006)
K.B. Quest, J. Geophys. Res. **93**, 9649 (1988)
K.B. Quest, in *AIP Conf. Ser.*, ed. by G.P. Zank, T.K. Gaisser. Particle Acceleration in Cosmic Plasmas, vol. 264, p. 349 (1992)
P. Richter, F.B.S. Paerels, J.S. Kaastra, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9325-4
C.T. Russell, in *AIP Conf. Ser.*, ed. by G. Li, G.P. Zank, C.T. Russell. The Physics of Collisionless Shocks, vol. 781, p. 3 (2005)
D. Ryu, H. Kang, E. Hallman, T.W. Jones, Astrophys. J. **593**, 599 (2003)
R.Z. Sagdeev, Rev. Plasma Phys. **4**, 23 (1966)
S. Schindler, W. Kapferer, W. Domainko et al., Astron. Astrophys. **435**, L25 (2005)
H. Schmitz, S.C. Chapman, R.O. Dendy, Astrophys. J. **579**, 327 (2002)
L. Spitzer, *Physics of Fully Ionized Gases*, 2nd edn. (Interscience, New York, 1962)
P. Tozzi, C. Norman, Astrophys. J. **546**, 63 (2001)
B.T. Tsurutani, R.P. Lin, J. Geophys. Res. **90**, 1 (1985)
A. Vladimirov, D.C. Ellison, A. Bykov, Astrophys. J. **652**, 1246 (2006)
G.M. Voit, M.L. Balogh, R.G. Bower, C.G. Lacey, G.L. Bryan, Astrophys. J. **593**, 272 (2003)
M.C. Weisskopf, J.P. Hughes, Astrophys. Update **2**, 55 (2006)
D. Winske, N. Omidi, J. Geophys. Res. **101**, 17287 (1996)
Y.B. Zel'dovich, Y.P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Academic, New York, 1967)

Chapter 8
Equilibration Processes in the Warm-Hot Intergalactic Medium

A.M. Bykov · F.B.S. Paerels · V. Petrosian

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9309-4

Abstract The Warm-Hot Intergalactic Medium (WHIM) is thought to contribute about 40–50% to the baryonic budget at the present evolution stage of the universe. The observed large scale structure is likely to be due to gravitational growth of density fluctuations in the post-inflation era. The evolving cosmic web is governed by non-linear gravitational growth of the initially weak density fluctuations in the dark energy dominated cosmology. Non-linear structure formation, accretion and merging processes, star forming and AGN activity produce gas shocks in the WHIM. Shock waves are converting a fraction of the gravitation power to thermal and non-thermal emission of baryonic/leptonic matter. They provide the most likely way to power the luminous matter in the WHIM. The plasma shocks in the WHIM are expected to be collisionless. Collisionless shocks produce a highly non-equilibrium state with anisotropic temperatures and a large differences in ion and electron temperatures. We discuss the ion and electron heating by the collisionless shocks and then review the plasma processes responsible for the Coulomb equilibration and collisional ionisation equilibrium of oxygen ions in the WHIM. MHD-turbulence produced by the strong collisionless shocks could provide a sizeable non-thermal contribution to the observed Doppler parameter of the UV line spectra of the WHIM.

Keywords Intergalactic medium · Shock waves · Galaxies: clusters: general

A.M. Bykov (✉)
A.F. Ioffe Institute of Physics and Technology, St. Petersburg 194021, Russia
e-mail: byk@astro.ioffe.ru

F.B.S. Paerels
Columbia Astrophysics Laboratory, Department of Astronomy, Columbia University,
550 West 120th Street, New York, NY 10027, USA
e-mail: frits@astro.columbia.edu

V. Petrosian
Department of Physics, Stanford University, Stanford, CA 94305, USA
e-mail: vahe@astronomy.stanford.edu

1 Introduction

Cosmological simulations of the large-scale structure (LSS) predict that about 40–50% of baryons at epoch $z < 2$ could reside in the Warm-Hot Intergalactic Medium (WHIM) with temperatures 10^5–10^7 K at moderate overdensities $\delta < 100$ (Cen and Ostriker 1999; Davé et al. 2001; Fang et al. 2002). The WHIM heating is due to shocks driven by gravitationally accelerated flows in the LSS structure formation scenario (e.g. Kang et al. 2007). Numerical simulations predict the observational signatures of the web gas as a function of redshift. The simulations account for feedback interactions between galaxies and the intergalactic medium, and demonstrate that the X-ray and ultraviolet O VI, O VII and O VIII lines and the H I Lyman alpha line are good tracers of low-density cosmic web filamentary structures (e.g. Hellsten et al. 1998; Tripp et al. 2000; Cen et al. 2001; Furlanetto et al. 2004). Intervening metal absorption systems of highly ionised C, N, O, Ne in the soft X-ray spectra of bright Active Galactic Nuclei (AGN) were suggested to be tracer of the WHIM. The predicted distribution of ion column densities in the WHIM absorbers is steep enough to provide only a few systems with $N_{\rm O\,VII} > 10^{15}$ cm^{-2} along an arbitrary chosen line of sight (e.g. Fang et al. 2002). Therefore, the detection of the WHIM is particularly difficult and requires very sensitive UV and X-ray detectors, both for absorption and for emission processes (e.g. Lehner et al. 2007; Nicastro et al. 2005; Kaastra et al. 2006; Takei et al. 2007 and Richter et al. 2008; Durret et al. 2008—Chaps. 3 and 4, this issue). Future projects and namely *Cosmic Origin Spectrograph* (*COS*), the *X-Ray Evolving Universe Spectrometer* (*XEUS*), *Constellation-X* and the *Diffuse Intergalactic Oxygen Surveyor* (*DIOS*) will increase the signal-to-noise ratio in the spectra allowing to study weak systems with $N_{\rm H\,I} < 10^{12.5}$ cm^{-2} and $N_{\rm O\,VII} < 10^{15}$ cm^{-2}. Simulations of spectra of the broad Lyα absorption lines and the highly ionised oxygen lines in the weak systems require thorough modelling of physical condition in the plasma (see e.g. Mewe 1990; Paerels and Kahn 2003; Kawahara et al. 2006). We discuss in this paper the heating and equilibration processes in the shocked WHIM plasma affecting the spectral simulations. A discussion of collisionless shock physics relevant to cosmological shocks in the WHIM can be found in Bykov et al. 2008—Chap. 7, this issue. To this end, in this paper we discuss first some specific features of collisionless shock heating of ions of different charge states providing highly non-equilibrium initial states just behind the magnetic ramp region that relaxes to an equilibrium state through Coulomb collisions, and the relation of the processes to simulations of emission/absorption spectra of the WHIM and observational data analysis.

2 WHIM Heating and Ion Temperature Evolution

The plasma ion heating in the WHIM is most likely due to cosmological shocks. The Alfvén Mach number of a shock propagating through an ionised gas of local overdensity $\delta = \rho/\langle\rho\rangle$ at the epoch z in the standard ΛCDM cosmology is determined by

$$\mathcal{M}_{\rm a} = v_{\rm sh}(4\pi\rho_{\rm i})^{1/2}/B \approx 20.6\, v_{\rm s7}\delta^{1/2}(1+z)^{3/2}(\Omega_{\rm b}h^2/0.02)^{1/2}B_{-9}^{-1}, \tag{1}$$

where B_{-9} is the magnetic field just before the shock, measured in nG and $v_{\rm s7}$ is the shock velocity in 10^7 cm s^{-1}, and $\langle\rho\rangle$ is the average density in the Universe.

The sonic Mach number for a shock propagating in a plasma of standard cosmic abundance is

$$\mathcal{M}_{\rm s} \approx 8.5\, v_{\rm s7}\,[T_4(1+f_{\rm ei})]^{-1/2}, \tag{2}$$

where T_4 is the plasma ion temperature measured in 10^4 K (typical for a preshock photoionised plasma) and $f_{\rm ei} = T_{\rm e}/T_{\rm i}$. An important plasma parameter is

$$\beta = \mathcal{M}_{\rm a}^2/\mathcal{M}_{\rm s}^2 \approx 6\delta(1+z)^3(\Omega_{\rm b}h^2/0.02)B_{-9}^{-2}[T_4(1+f_{\rm ei})].$$

It is the ratio of the thermal and magnetic pressures. In hot X-ray clusters of galaxies the beta parameter is ∼100 for ∼μG magnetic fields in the clusters. The most uncertain parameter is the magnetic field value in the WHIM allowing for both $\beta \sim 1$ and $\beta \gg 1$ cases.

In a supercritical collisionless shock the conversion of kinetic energy of an initially cold flow to the ion distribution with high kinetic temperature occurs in the thin ion viscous jump. The width of the ion viscous jump $\Delta_{\rm vi}$ in a collisionless shock propagating through a plasma with $\beta \sim 1$ is typically of the order of a ten to a hundred times of the ion inertial length $l_{\rm i}$ defined as $l_{\rm i} = c/\omega_{\rm pi} \approx 2.3 \times 10^7 n^{-0.5}$ cm. Here $\omega_{\rm pi}$ is the ion plasma frequency. The ion inertial length in the WHIM can be estimated as

$$l_{\rm i} \approx 5.1 \times 10^{10}\delta^{-1/2}(1+z)^{-3/2}(\Omega_{\rm b}h^2/0.02)^{-1/2}\ {\rm cm}.$$

The width of the shock transition region for magnetic field is also $\gtrsim 10 l_{\rm i}$ for a quasi-perpendicular shock, but it is often about ten times wider for quasi-parallel shocks.

Properties of nonrelativistic shocks in a hot, low magnetised plasma with high $\beta \gg 1$ are yet poorly studied. Measurements from the *ISEE 1* and *ISEE 2* spacecrafts were used by Farris et al. (1992) to examine the terrestrial bow shock under high beta conditions. These measurements were compared with and found to be in agreement with the predicted values of the Rankine-Hugoniot relations using the simple adiabatic approximation and a ratio of specific heats, γ, of 5/3. Large magnetic field and density fluctuations were observed, but average downstream plasma conditions well away from the shock were relatively steady, near the predicted Rankine-Hugoniot values. The magnetic disturbances persisted well downstream and a hot, dense ion beam was detected leaking from the downstream region of the shock. The observation proved the existence of collisionless shocks in high beta plasma, but a detailed study of high beta shock structure is needed for cosmological plasmas.

We discuss in the next section the ion heating in collisionless shocks illustrating the most important features of the process with the results of a hybrid simulation of the oxygen ions heating in a quasi-perpendicular shock considered earlier by Bykov et al. 2008—Chap. 7, this issue.

2.1 Collisionless Shock Heating of the Ions

Ion heating mechanisms in collisionless shocks depend on the shock Alfvén Mach number, the magnetic field inclination angle ($\theta_{\rm n}$), plasma parameter β and the composition of the incoming plasma flow. The structure of a supercritical shock is governed by the ion flows instabilities (see e.g. Kennel et al. 1985; Lembege et al. 2004; Burgess et al. 2005). In a quasi-parallel shock ($\theta_{\rm n} \lesssim 45°$) a mixed effect of a sizeable backstreaming ion fraction and the ions scattered by the strong magnetic field fluctuations (filling the wide shock transition region) results in the heating of ions in the downstream region. The ions reflected and slowed down by an electric potential jump $\delta\phi$ at the shock ramp of a quasi-perpendicular ($\theta_{\rm n} \gtrsim 45°$) shock constitute a multi-stream distribution just behind a relatively thin magnetic ramp as it is seen in Figs. 1 and 2 (left panel). The O VII phase densities and distribution functions were simulated with a hybrid code for a quasi-perpendicular ($\theta_{\rm n} = 80°$) shock in a hydrogen-helium dominated plasma (see Bykov et al. 2008—Chap. 7, this issue). Phase densities $x - v_x$, $x - v_y$, $x - v_z$ of the O VII ion are shown in Fig. 1 in the reference frame

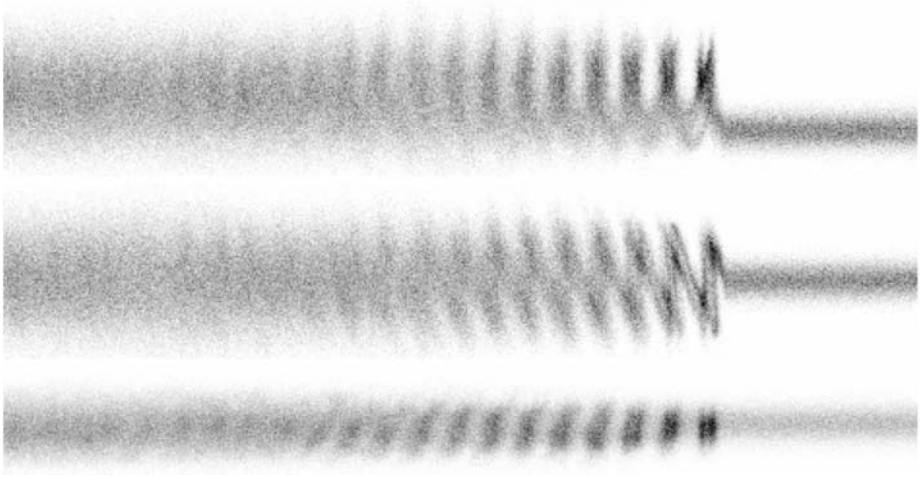


Fig. 1 O VII phase density in a hybrid simulated quasi-perpendicular shock (80° inclination). The shock is moving from *left* to *right* in the reference frame where the particle reflecting wall is at rest. The figures show the oxygen phase densities in x–v_x, x–v_y and x–v_z projections from *top* to *bottom* respectively. The size of the simulation box in x-dimension is about 300 l_i

where the particle reflecting wall (at far right) is at rest and the shock is moving. The shock is propagating along the x-axis from the left to the right and the magnetic field is in the x–z plane. The system is periodic in the y dimension. The incoming plasma beam in the simulation was composed of protons (90%), alpha particles (9.9%) and a dynamically insignificant fraction of oxygen ions (O VII) with the upstream plasma parameter $\beta \sim 1$. The ions do not change their initial charge states in a few gyro-periods while crossing the cosmological shock ramp where the Coulomb interactions are negligible.

The simulated data in Fig. 1 show the ion velocities phase mixing resulting in a thermal-like broad ion distribution at a distance of some hundreds of ion inertial lengths in the shock downstream (see the right panel in Fig. 2). It is also clear in Fig. 1 that the shocked ion distribution tends to have anisotropy of the effective temperature. The temperature anisotropy $T_\perp \sim 3T_\parallel$ relative to the magnetic field was found in that simulation. Moreover, the hybrid simulation shows that the $T_\perp$ of the O VII is about 25 times higher than the effective perpendicular temperature of the protons. Thus the ion downstream temperature declines from the linear dependence on the ion mass. The simulations show excessive heating of heavy ions in comparison with protons.

Lee and Wu (2000) proposed a simplified analytical model to estimate the ion perpendicular temperature dependence on Z/A, where $m_\mathrm{i} = Am_\mathrm{p}$. Specifically, the model predicts the ratio of the ion gyration velocity v_ig2 in the downstream of a perpendicular shock ($\theta_\mathrm{n} \sim 90°$) to the velocity of the incident ion in the shock upstream, v_1,

$$\frac{v_\mathrm{ig2}}{v_1} = \left| \left(1 - \alpha \frac{Z}{A}\right)^{1/2} - \frac{B_\mathrm{t1}}{B_\mathrm{t2}} \right|, \tag{3}$$

where $\alpha = 2\mathrm{e}\,\delta\phi/m_\mathrm{p}v_1^2 < 1$, and the potential jump $\delta\phi$ is calculated in the shock normal frame (see Lee and Wu 2000). The model is valid for the ions with gyroradii larger than the shock transition width Δ_vi. It is not a fair approximation for the protons, but it is much better for temperature estimation of heavy ions just behind the shock magnetic ramp. The model

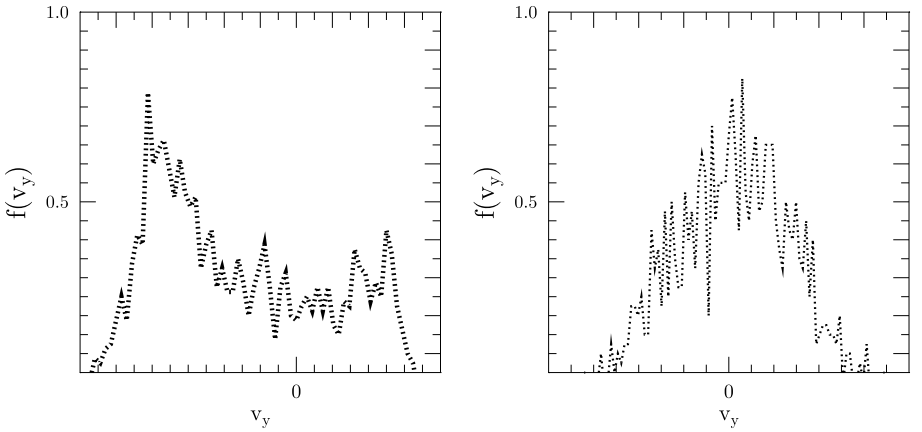


Fig. 2 Hybrid simulated O VII distribution function (normalised) as a function of a random velocity component $\delta v_y = v_y - \langle v_y \rangle$ transverse to the downstream magnetic field in a quasi-perpendicular shock (80° inclination). The shock propagates along the x-axis, while the initial regular magnetic field is in the x–z plane. In the *left panel* the distribution in the viscous velocity jump is shown. The *right panel* shows the distribution behind the jump at the position of the left end in Fig. 1. Multi-velocity structure of the flow is clearly seen in the *left panel*, while it is relaxing to quasi-Maxwellian in the *right panel*

of ion heating in the fast, supercritical quasi-perpendicular ($\theta_n \gtrsim 45°$) shocks of $\mathcal{M}_a \gtrsim 3$ predicts a higher downstream perpendicular temperature for the ions with larger A/Z.

2.2 Collisionless Heating of the Electrons

The initial electron temperature just behind the viscous ion jump of a cosmological shock depends on the collisionless heating of the electrons. The only direct measurements of the electron heating by collisionless shocks are those in the Heliosphere. The interplanetary shock data compiled by Schwartz et al. (1988) show a modest, though systematic departure of the electron heating from that which would result from the approximately constant ratio of the perpendicular temperature to the magnetic field strength (i.e. adiabatic heating). Thus, some modest non-adiabatic electron collisionless heating is likely present. In the case of a nonradiative supernova shock propagating through *partially ionised* interstellar medium the ratio T_e/T_i in a thin layer (typically $< 10^{17}$ cm) just behind a shock can be tested using the optical diagnostics of broad and narrow Balmer lines (e.g. Raymond 2001). High resolution *Hubble Space Telescope (HST)* Supernova remnant (SNR) images make that approach rather attractive. A simple scaling $T_e/T_i \propto v_{sh}^{-2}$ was suggested by Ghavamian et al. (2007) to be consistent with the optical observations of SNRs.

Strong shocks are thought to transfer a sizeable fraction of the bulk kinetic energy of the flow into large amplitude nonlinear waves in the magnetic ramp region. The thermal electron velocities in the ambient medium are higher than the shock speed if the shock Mach number $\mathcal{M}_s < \sqrt{m_p/m_e}$, allowing for a nearly-isotropic angular distribution of the electrons. Non-resonant interactions of these electrons with large-amplitude turbulent fluctuations in the shock transition region could result in collisionless heating and pre-acceleration of the electrons (Bykov and Uvarov 1999; Bykov 2005). They calculated the electron energy spectrum in the vicinity of the shock waves and showed that the heating and pre-acceleration of the electrons occur on a scale of the order of several hundred ion inertial lengths in the vicinity of the viscous discontinuity. Although the electron distribution function is in a significantly

non-equilibrium state near the shock front, its low energy part can be approximated by a Maxwellian distribution. The effective electron temperature just behind the front, obtained in this manner, increases with the shock wave velocity as $T_e \propto v_{sh}^b$ with $b \leq 2$. They also showed that if the electron transport in the shock transition region is due to turbulent advection by strong vortex fluctuations of the scale of about the ion inertial length, then the nonresonant electron heating is rather slow (i.e. $b \leq 0.5$). The highly developed vortex-type turbulence is expected to be present in the transition regions of very strong shocks. That would imply that the initial $T_e/T_i \propto v_{sh}^{(b-a)}$ just behind the transition region would decrease with the shock velocity for $\mathcal{M}_s \gg 1$. Here the index a is defined by the relation $T_i \propto v_{sh}^a$ for a strong shock. The degree of electron-ion equilibration in a collisionless shock is a declining function of shock speed. In the case of strong vortex-type turbulence in the shock transition region one expects in the standard ion heating case with $a = 2$ and rather small b to have $(a - b) \lesssim 2$. That T_e/T_i scaling is somewhat flatter, but roughly consistent with, that advocated by Ghavamian et al. (2007). On the other hand in a collisionless shock of a moderate strength $\mathcal{M}_s < 10$ the electron transport through the magnetic ramp region could be diffusive, rather than by the turbulent advection by a strong vortexes. That results in a larger degree of the collisionless electron heating/equilibration in the shocks as it is shown in Fig. 4 of the paper by Bykov and Uvarov (1999). Recently, Markevitch and Vikhlinin (2007) argued for the collisionless heating/equilibration of the electron temperature in the bow shock of $\mathcal{M}_s \sim 3$ in the 1E 0657–56 cluster.

If the local Mach number $\mathcal{M}_s$ of the incoming flow in a strong shock wave exceeds $\sqrt{m_p/m_e}$, which could occur in the cluster accretion shocks, the thermal electron distribution becomes highly anisotropic and high frequency whistler type mode generation effects could become important. Levinson (1996) performed a detailed study of resonant electron acceleration by the whistler mode for fast MHD shock waves. Electron heating and Coulomb relaxation in the strong accretion shocks in clusters of galaxies was discussed in details by Fox and Loeb (1997).

We summarise this section concluding that a collisionless shock produces in the downstream flow a highly non-equilibrium plasma state with strongly different temperatures of the electrons and ions of different species. Moreover, the postshock ion temperatures are anisotropic. The width of the collisionless shock transition region is smaller by many orders of magnitude than the Coulomb mean free path (that is of a kiloparsec range). We consider now the structure and the processes in the postshock Coulomb equilibration layers in the WHIM.

3 Coulomb Relaxation of Temperatures in the WHIM

3.1 Relaxation of the Ion Temperature Anisotropy

A plasma flow partly randomised just behind the viscous ion jump in a collisionless shock transition has an anisotropic velocity distribution with respect to the mean magnetic field. In Fig. 1 the oxygen phase densities in the v_x–x and v_y–x projections transverse to the mean field have a substantially wider distribution than that in the projection v_z–x parallel to the magnetic field. In many cases the ion velocity distributions (like that in Fig. 2 right) can be approximated with a quasi-Maxwellian distribution introducing some effective kinetic temperature (more exactly it is the second moment of the distribution). Then the parallel and perpendicular (to the mean field) temperatures are different and we approximate the 3D

particle distribution as

$$f(v_\perp, v_\parallel) = \left(\frac{m}{2\pi T_\perp}\right)\left(\frac{m}{2\pi T_\parallel}\right)^{1/2} \exp\left(-\frac{m v_\perp^2}{2T_\perp} - \frac{m v_\parallel^2}{2T_\parallel}\right). \tag{4}$$

We will measure the temperatures in energy units in most of the equations below (thus $k_B T \to T$). Ichimaru and Rosenbluth (1970) (see also a comment by Kaiser 1979) obtained the following equation to describe the ion anisotropy relaxation

$$\frac{dT_\perp}{dt} = -\frac{1}{2}\frac{dT_\parallel}{dt} = -\frac{T_\perp - T_\parallel}{\tau_i}. \tag{5}$$

Here

$$\tau_i^{-1} = \frac{8\pi^{1/2} n_i Z^2 e^4 \ln\Lambda}{15 m_i^{1/2} T_{eff}^{3/2}}, \tag{6}$$

where $\ln\Lambda$ is the Coulomb logarithm and the effective ion temperature is defined as

$$\frac{1}{T_{eff}^{3/2}} = \frac{15}{4}\int_{-1}^{1} d\mu \frac{\mu^2(1-\mu^2)}{[(1-\mu^2)T_\perp + \mu^2 T_\parallel]^{3/2}}. \tag{7}$$

These equations were obtained under the assumption that the electrons have no dynamical role, but provide a static dielectric background to the ions. We can directly apply (5–7) to isotropisation of the plasma field particles (i.e. protons in our case). Isotropisation of the minor ion components is mainly due to their interactions with protons and helium because of a low metal number density in cosmic plasma.

3.2 Relaxation of the Ion and Electron Temperatures

Following shock heating and the temperature isotropisation a quasi-Maxwellian distribution will be established within all plasma components, i.e. groups of identical particles, after a time scale given by (6). The effective temperatures differs strongly between the components. All the plasma particles will undergo Coulomb collisions with the protons, alpha particles and electrons dominating the WHIM plasma, resulting in the temperature equilibration. Spitzer (1962) found that the temperature relaxation of a test particle of type a with plasma field particles can be approximately estimated from

$$\frac{dT_a}{dt} = \frac{T_p - T_a}{\tau_{ap}} \tag{8}$$

where

$$\tau_{ap}^{-1} = \frac{8(2\pi)^{1/2}}{3}\frac{n_p Z^2 e^4 \ln\Lambda}{m_a m_p}\left(\frac{T_p}{m_p} + \frac{T_a}{m_a}\right)^{-3/2}. \tag{9}$$

In the thermal equilibrium state the postshock plasma must have a single equilibrium temperature T_{eq}. In a fully ionised plasma without energy exchange with external components (i.e. radiation or plasma wave dissipation) T_{eq} can be found from the condition of constant pressure in the plane shock downstream resulting in

$$T_{eq} = \sum n_a T_{a0} / \sum n_a. \tag{10}$$

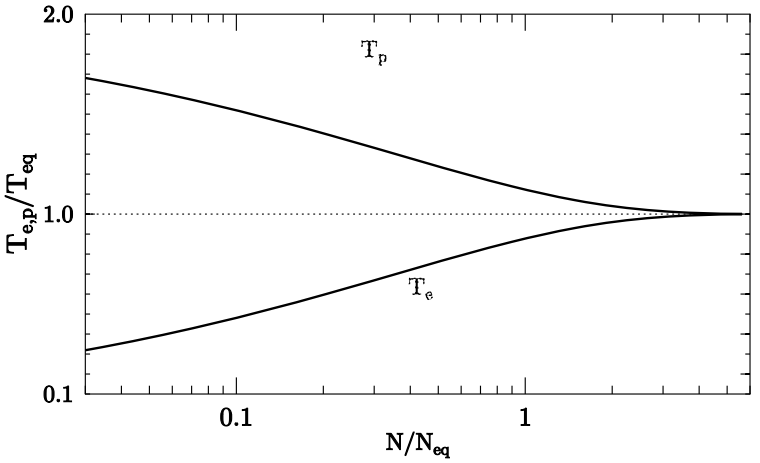


Fig. 3 Postshock temperature equilibration between the ion and electron components due to the Coulomb interactions

In cosmic plasmas it is often a fair approximation to estimate $T_{\rm eq}$ from the equation $2\,T_{\rm eq} = T_{\rm e} + T_{\rm p}$. Then following Sivukhin (1966) the charged particle equilibration can be approximately described through the equation

$$\ln\left|\frac{\sqrt{T_{\rm e}}-\sqrt{T_{\rm eq}}}{\sqrt{T_{\rm e}}+\sqrt{T_{\rm eq}}}\right| = -\frac{t}{\tau_{\rm eq}} - \frac{2}{3}\left(\frac{T_{\rm e}}{T_{\rm eq}}\right)^{3/2} - 2\left(\frac{T_{\rm e}}{T_{\rm eq}}\right)^{1/2} + C_{\rm e}, \tag{11}$$

where $C_{\rm e}$ is a constant to be determined from the initial temperature $T_{\rm a0}$ of a relaxing component $a = {\rm e},\, p$,

$$\tau_{\rm eq}^{-1} \approx \frac{16(2\pi)^{1/2}}{3}\frac{n_{\rm p}e^4\ln\Lambda}{m_{\rm e}m_{\rm p}}\left(\frac{T_{\rm eq}}{m_{\rm e}}\right)^{-3/2}. \tag{12}$$

Equations (11) and (12) allow to calculate the structure of relaxation layers to be seen behind a collisionless shock in the WHIM. In Fig. 3 we illustrate the postshock equilibration of initially cold electrons with the protons initially heated at the ion viscous jump of a collisionless shock transition. The width of the ion viscous jump in cosmological shocks is negligible compared with the equilibration length $x_{\rm eq} = u_2\tau_{\rm eq}$, where u_2 is the downstream flow velocity in the shock rest frame. The characteristic column density $N_{\rm eq} = n_2 x_{\rm eq}$ to be traversed by protons and electrons in the downstream plasma (of a number density n_2) before the temperature equilibration, as it follows from (12), does not depend on the plasma number density, and $N_{\rm eq} \propto v_{\rm sh}^4$. The corresponding shocked WHIM column density can be expressed through the shock velocity v_7 measured in 100 km s^{-1}, assuming a strong shock where $u_2 = v_{\rm sh}/4$:

$$N_{\rm eq} \approx 5\times 10^{17} v_7^4 [\ln\Lambda]^{-1}\ {\rm cm}^{-2},$$

or through the shocked WHIM gas temperature T_6 (measured in 10^6 K):

$$N_{\rm eq} \approx 2.5\times 10^{17} T_6^2 [\ln\Lambda]^{-1}\ {\rm cm}^{-2}.$$

The Coulomb logarithm for the WHIM condition is $\ln \Lambda \sim 40$. It follows from Fig. 3 that in the Coulomb relaxation model the postshock plasma column density $N_{\rm H} > 3N_{\rm eq}$ ensures the equilibration at a level better than 1%.

The metal ions can be initially heated at the shock magnetic ramp to high enough temperatures $> AT_{\rm p}$ (see e.g. Korreck et al. (2007) for a recent analysis of interplanetary collisionless shock observations with *Advanced Composition Explorer*). However, in a typical case the depth $N > 3N_{\rm eq}$ is enough for the ion temperature equilibration.

To study UV and X-ray spectra of the weak systems ($N_{\rm H\,I} < 10^{12.5}$ cm^{-2} and $N_{\rm O\,VII} < 10^{15}$ cm^{-2}) modelling of shocked filaments of $N_{\rm H} \lesssim 10^{17}$ cm^{-2} would require an account of non-equilibrium effects of low electron temperature $T_{\rm e}/T_{\rm eq} < 1$.

3.3 Effect of Postshock Plasma Micro-Turbulence on the Line Widths

We consider above only the WHIM temperature evolution due to the Coulomb equilibration processes. Shocks producing the WHIM could propagate through inhomogeneous (e.g. clumpy) matter. The interaction of a shock with the density inhomogeneities results in the generation of MHD-waves (Alfvén Mach and magnetosonic) in the shock downstream (see e.g. Vainshtein et al. 1993). The MHD-wave dissipation in the shock downstream could selectively heat ions, being a cause of a non-equilibrium $T_{\rm e}/T_{\rm i}$ ratio. In case of a strong collisionless shock propagating in a turbulent medium, cosmic ray acceleration could generate a spectrum of strong MHD-fluctuations (see e.g. Blandford and Eichler 1987; Bell 2004; Vladimirov et al. 2006). These MHD-fluctuations could carry a substantial fraction of the shock ram pressure to the upstream and then to the downstream providing a heating source throughout the downstream. The velocity fluctuations could also produce non-thermal broadening of the lines. The amplitude of bulk velocity fluctuations is about the Alfvén Mach velocity since $\delta B \sim B$ in the shock precursor. The Doppler parameter b derived from high resolution UV spectra of the WHIM (see e.g. Lehner et al. 2007; Richter et al. 2008—Chap. 3, this issue):

$$b^2 = \frac{2T}{m_{\rm a}} + b_{\rm nt}^2$$

would have in the micro-turbulent limit a non-thermal contribution

$$b_{\rm nt}^2 = \frac{2v_{\rm turb}^2}{3} = \frac{C_\nu B^2}{6\pi \langle \rho \rangle \delta}.$$

The factor C_ν here accounts for the amplitude and spectral shape of the turbulence. For a strong MHD-turbulence that was found in the recent models of strong collisionless shocks with efficient particle acceleration (e.g. Vladimirov et al. 2006) one can get $C_\nu \sim 1$ in the WHIM, and then

$$b_{\rm nt} \approx 4B_{-9}(\Omega_{\rm b}h^2/0.02)^{-1/2}(1+z)^{-3/2}\delta^{-1/2}\ {\rm km\,s^{-1}}.$$

The estimation of $b_{\rm nt}$ given above for a strong Alfvén Mach turbulence may be regarded as an upper limit for a system with a modest MHD turbulence. The non-thermal Doppler parameter does not depend on the ion mass, but $b_{\rm nt} \propto \delta^{-1/2}B(z)$. Thus the account of $b_{\rm nt}$ could be important for high resolution spectroscopy of metal lines, especially if the strong shocks can indeed highly amplify local magnetic fields. That is still to be confirmed, but a recent high resolution observation of a strong Balmer-dominated shock on the eastern side

of Tycho's supernova remnant with the *Subaru Telescope* supports the existence of a cosmic ray shock precursor where gas is heated and accelerated ahead of the shock (Lee et al. 2007). High resolution UV spectroscopy of the WHIM could allow to constrain the intergalactic magnetic field. Internal shocks in hot X-ray clusters of galaxies have modest Mach numbers and the effect of the Alfvén Mach turbulence is likely less prominent than that in the strong accretion shocks in clusters and in the cosmic web filaments. X-ray line broadening by large scale bulk motions in the hot intracluster medium was discussed in detail by Fox and Loeb (1997) and Inogamov and Sunyaev (2003).

4 Ionisation State of the WHIM

To simulate absorption spectra of bright quasars in the intervening WHIM filaments (e.g. Kawahara et al. 2006 as an example of such a modelling) one should solve the ionisation balance equations for the charge states of metal ions with account taken of all the LTE processes and also the nonthermal particle contribution (see Porquet et al. 2001 for a discussion of a role of non-relativistic super-thermal distributions). The ionisation balance equation can be written as

$$\dot{n}_{\rm q} = n_{\rm e}\left[n_{\rm q-1}C_{\rm q-1} - n_{\rm q}C_{\rm q} - n_{\rm q}\alpha_{\rm q} + n_{\rm q+1}\alpha_{\rm q+1}\right] + \sum_{j={\rm H,\,He,\,He^+}} n_{\rm j}\left[n_{\rm q-1}V^{\rm ion}_{\rm j,q-1} - n_{\rm q}(V^{\rm rec}_{\rm j,q} + V^{\rm ion}_{\rm j,q}) + n_{\rm q+1}V^{\rm rec}_{\rm j,q+1}\right] + n_{\rm q-1}R_{\rm q-1} - n_{\rm q}R_{\rm q}.$$

Here q is the charge state of an element, $C_{\rm q}$ is the collisional (and autoionisation) rate $q \to q+1$, $\alpha_{\rm q}$ are the radiative and the dielectronic ionisation rates $q \to q-1$ (in cm^3 s^{-1}), $V^{\rm rec}_{\rm j,q}$. $V^{\rm ion}_{\rm j,q}$ are the charge exchange rates with the ion j (in cm^3 s^{-1}) and $R_{\rm q}$ is the photoionisation rate of an ion (in s^{-1}). The rates of different processes can be calculated for different temperature regimes (see Kaastra et al. 2008—Chap. 9, this issue). We just limit our discussion here to one example of such a simulation.

In Fig. 4 we illustrate *collisional ionisation equilibrium* curves of oxygen ions in the present epoch (at $z = 0$) as a function of WHIM density δ. The various boundaries separate regimes under which a certain process does, or does not attain equilibrium over a Hubble time. We show in Fig. 4 some critical boundaries for kinetic and thermal equilibrium. The solid line in the upper left hand corner labelled '$t_{\rm ei} < t_{\rm H}$' indicates where electron and proton fluids reach kinetic equilibrium (proton temperature equal to electron temperature) in a Hubble time ($t_{\rm H}$): at low density and high temperature, such equilibrium does not attain. Low density gas will not radiatively cool over a Hubble time to the left of the boundaries marked '$t_{\rm cool} > t_{\rm H}$'. The cooling time was calculated for collisionally ionised gas. The solid curve labelled '*adiabatic evolution*' indicates the locus of gas that has only undergone adiabatic compression or expansion since high redshift (initial condition $T_{\rm e} \sim 10^4$ K); all shock-heated gas will be above this line right after passing through a shock. The shaded area in the upper right hand corner indicates the regime where the collisional ionisation timescale is shorter than the photoionisation timescale, for ionisation O VIII → O IX. The two boundaries labelled '$t_{\rm coll} < t_{\rm H}$' indicate where the collisional ionisation timescale becomes shorter than the Hubble time. At lower temperature, the ionisation balance cannot be in (collisional) equilibrium. The steep solid curve labelled '$t_{\rm rec}$(Oxygen) $> t_{\rm H}$' indicates where the radiative recombination timescale (O IX → O VIII) exceeds the Hubble time since there is no recombination at low densities.

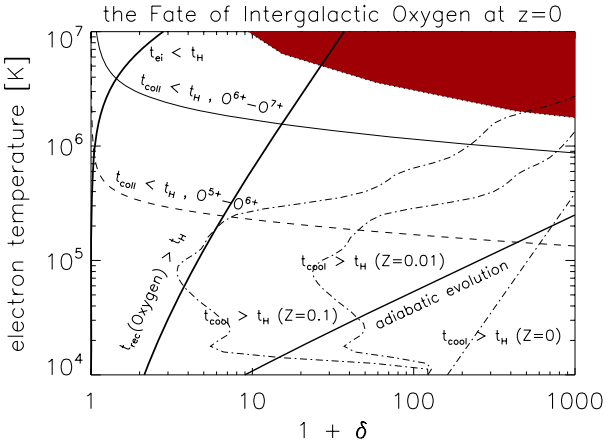


Fig. 4 Phase diagram for oxygen in the IGM at redshift $z = 0$. Density is parameterised by the δ value as was defined above. The *solid line* in the *upper left* hand corner labelled '$t_{\rm ei} < t_H$' indicates where electron and proton fluids reach equilibrium in a Hubble time ($t_{\rm H}$). Low density gas will not radiatively cool over a Hubble time to the *left* of the boundaries marked '$t_{\rm cool} > t_{\rm H}$'. The *three curves* are labelled with the metallicity, Z, expressed as a fraction of Solar metallicity. The *solid curve* labelled '*adiabatic evolution*' indicates the locus of gas that has only undergone adiabatic compression or expansion since high redshift (initial condition $T_{\rm e} \sim 10^4$ K). Critical boundaries for the ionisation equilibrium of oxygen are: the *shaded area* in the *upper right* hand corner indicates the regime where the collisional ionisation timescale is shorter than the photoionisation timescale, for ionisation O VIII → O IX. The two boundaries labelled '$t_{\rm coll} < t_{\rm H}$' indicate where the collisional ionisation timescale becomes shorter than the Hubble time. At lower temperature, the ionisation balance cannot be in (collisional) equilibrium. *Upper* (*solid*) *curve* is for ionisation O VII (O^{+6}) → O VIII (O^{+7}), *lower* (*dashed*) curve for O VI (O^{+5}) → O VII (O^{+6}). The *steep solid curve* labelled '$t_{\rm rec}$(Oxygen) > $t_{\rm H}$' indicates where the radiative recombination timescale (O IX → O VIII) exceeds the Hubble time (no recombination at low densities)

Note that we illustrate here only the collisional equilibrium case. For the more appropriate case of radiative cooling in photoionisation equilibrium, the cooling times will be even longer, due to the fact that a photoionised plasma is highly overionised compared to the characteristic ionisation- and excitation potentials, which suppresses the (very effective) collisional cooling contribution. In our figure, the lines $t_{\rm cool} = t_{\rm H}$ will shift to the right if we calculate with the probably more realistic case of photoionisation equilibrium. However, as we argued above the collisional equilibrium is also of interest, since it represents a conservative case. The range of ionisation states of oxygen in the WHIM filaments can be observed in the absorption spectra of bright quasars.

5 Conclusions

We discussed some specific features of the WHIM heating processes by collisionless plasma shocks driven by gravitationally accelerated flows in the LSS structure formation scenario.

- Collisionless shock heating of ions in the WHIM results in a highly non-equilibrium initial state with a strongly anisotropic quasi-Maxwellian ion distribution just behind a very thin magnetic ramp region. The ion temperatures just behind the shock depend on the ion atomic weight, the charge state and the shock magnetic field inclination. The ion temperatures could decline from a simple linear scaling with the ion mass providing an excessive heating of heavy ions.

- The ion and electron temperatures relax to the equilibrium state through Coulomb collisions in a layer of the depth $N_{\rm eq} \sim 10^{16} v_7^4$ cm^{-2} behind a shock of velocity v_7.
- Strong collisionless plasma shocks with an efficient Fermi acceleration of energetic particles can generate strong MHD waves in the upstream and downstream regions that will result in a non-thermal broadening of the emission/absorption lines in the observed WHIM spectra.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". A.M.B. thanks M. Yu. Gustov for his help with hybrid shock simulations, he acknowledges the RBRF grant 06-02-16844 and a support from RAS Presidium Programs. A support from NASA ATP (NNX07AG79G) is acknowledged.

References

A.R. Bell, Mon. Not. R. Astron. Soc. **353**, 550 (2004)
R. Blandford, D. Eichler, Phys. Rep. **154**, 1 (1987)
D. Burgess, E.A. Lucek, M. Scholer et al., Space Sci. Rev. **118**, 205 (2005)
A.M. Bykov, Adv. Space Res. **36**, 738 (2005)
A.M. Bykov, Y.A. Uvarov, J. Exp. Theor. Phys. **88**, 465 (1999)
A. Bykov, K. Dolag, F. Durret, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9312-9
R. Cen, J.P. Ostriker, Astrophys. J. **514**, 1 (1999)
R. Cen, T.M. Tripp, J.P. Ostriker, E.B. Jenkins, Astrophys. J. Lett. **559**, L5 (2001)
R. Davé, R. Cen, J.P. Ostriker et al., Astrophys. J. **552**, 473 (2001)
F. Durret, J.S. Kaastra, J. Nevalainen, T. Ohashi, N. Werner, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9313-8
T. Fang, G.L. Bryan, C.R. Canizares, Astrophys. J. **564**, 604 (2002)
M.H. Farris, C.T. Russell, M.F. Thomsen, J.T. Gosling, J. Geophys. Res. **97**, 19121 (1992)
D.C. Fox, A. Loeb, Astrophys. J. **491**, 459 (1997)
S.R. Furlanetto, J. Schaye, V. Springel, L. Hernquist, Astrophys. J. **606**, 221 (2004)
P. Ghavamian, J.M. Laming, C.E. Rakowski, Astrophys. J. Lett. **654**, L69 (2007)
U. Hellsten, N.Y. Gnedin, J. Miralda-Escudé, Astrophys. J. **509**, 56 (1998)
S. Ichimaru, M.N. Rosenbluth, Phys. Fluids **13**, 2778 (1970)
N.A. Inogamov, R.A. Sunyaev, Astron. Lett. **29**, 791 (2003)
J.S. Kaastra, N. Werner, J.W.A. den Herder et al., Astrophys. J. **652**, 189 (2006)
J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-y
T.B. Kaiser, Phys. Fluids **22**, 593 (1979)
H. Kang, D. Ryu, R. Cen, J.P. Ostriker, Astrophys. J. **669**, 729 (2007)
H. Kawahara, K. Yoshikawa, S. Sasaki et al., Publ. Astron. Soc. Jpn. **58**, 657 (2006)
C.F. Kennel, J.P. Edmiston, T. Hada, Geophys. Monograph. Ser. **34**, 1 (1985)
K.E. Korreck, T.H. Zurbuchen, S.T. Lepri, J.M. Raines, Astrophys. J. **659**, 773 (2007)
L.C. Lee, B.H. Wu, Astrophys. J. **535**, 1014 (2000)
J.-J. Lee, B.-C. Koo, J. Raymond et al., Astrophys. J. Lett. **659**, L133 (2007)
N. Lehner, B.D. Savage, P. Richter et al., Astrophys. J. **658**, 680 (2007)
B. Lembege, J. Giacalone, M. Scholer et al., Space Sci. Rev. **110**, 161 (2004)
A. Levinson, Mon. Not. R. Astron. Soc. **278**, 1018 (1996)
M. Markevitch, A. Vikhlinin, Phys. Rep. **443**, 1 (2007)
R. Mewe, in *Physical Processes in Hot Cosmic Plasmas*, ed. by W. Brinkmann, A.C. Fabian, F. Giovannelli (Kluwer, Dordrecht, 1990), p. 39
F. Nicastro, S. Mathur, M. Elvis et al., Nature **433**, 495 (2005)
F.B.S. Paerels, S.M. Kahn, Annu. Rev. Astron. Astrophys. **41**, 291 (2003)
D. Porquet, M. Arnaud, A. Decourchelle, Astron. Astrophys. **373**, 1110 (2001)
J.C. Raymond, Space Sci. Rev. **99**, 209 (2001)
P. Richter, F.B.S. Paerels, J.S. Kaastra, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9325-4
S.J. Schwartz, M.F. Thomsen, S.J. Bame, J. Stansberry, J. Geophys. Res. **93**, 12923 (1988)
D.V. Sivukhin, in *Reviews of Plasma Physics*, vol. 4, p. 93 (1966)

L. Spitzer, *Physics of Fully Ionized Gases*, 2nd edn. (Interscience, New York, 1962)
Y. Takei, J.P. Henry, A. Finoguenov et al., Astrophys. J. **655**, 831 (2007)
T.M. Tripp, B.D. Savage, E.B. Jenkins, Astrophys. J. Lett. **534**, L1 (2000)
S.I. Vainshtein, A.M. Bykov, I. Toptygin, *Turbulence, Current Sheets, and Shocks in Cosmic Plasma* (Gordon & Breach, New York, 1993)
A. Vladimirov, D.C. Ellison, A. Bykov, Astrophys. J. **652**, 1246 (2006)

Chapter 9
Thermal Radiation Processes

J.S. Kaastra · F.B.S. Paerels · F. Durret · S. Schindler · P. Richter

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9310-y

Abstract We discuss the different physical processes that are important to understand the thermal X-ray emission and absorption spectra of the diffuse gas in clusters of galaxies and the warm-hot intergalactic medium. The ionisation balance, line and continuum emission and absorption properties are reviewed and several practical examples are given that illustrate the most important diagnostic features in the X-ray spectra.

Keywords Atomic processes · Radiation mechanisms: thermal · Intergalactic medium · X-rays: general

1 Introduction

Thermal X-ray radiation is an important diagnostic tool for studying cosmic sources where high-energy processes are important. Examples are the hot corona of the Sun and of stars,

J.S. Kaastra (✉) · F.B.S. Paerels
SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands
e-mail: j.kaastra@sron.nl

J.S. Kaastra
Astronomical Institute, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, The Netherlands

F.B.S. Paerels
Department of Astronomy, Columbia University, New York, NY, USA

F. Durret
Institut d'Astrophysique de Paris, CNRS, UMR 7095, Université Pierre et Marie Curie, 98bis Bd Arago, 75014 Paris, France

S. Schindler
Institute for Astro- and Particle Physics, University of Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria

P. Richter
Physics Institute, Potsdam University, Am Neuen Palais 10, 14469 Potsdam, Germany

Solar and stellar flares, supernova remnants, cataclysmic variables, accretion disks in binary stars and around black holes (Galactic and extragalactic), the diffuse interstellar medium of our Galaxy or external galaxies, the outer parts of Active Galactic Nuclei (AGN), the hot intracluster medium, the diffuse intercluster medium. In all these cases there is thermal X-ray emission or absorption.

In the present paper we focus upon the properties of the X-ray spectrum in the diffuse gas within and between galaxies, clusters and the large scale structures of the Universe. As this gas has very low densities, the level populations in the atoms are not governed by Saha-like equations, but instead most of the atoms will be in or near the ground state. This simplifies the problem considerably. Furthermore, in these tenuous media we need to consider only gas with small or moderate optical depth; the radiative transport is therefore simple. Stellar coronae have similar physical conditions, but there the densities are higher so that in several cases the emergent spectrum has density-dependent features. Due to the low densities in our sources, these density effects can be ignored in most cases; however, for the lowest density gas photoionisation effects must be taken into account.

We show in this paper that it is possible to derive many different physical parameters from an X-ray spectrum: temperature, density, chemical abundances, plasma age, degree of ionisation, irradiating continuum, geometry etc.

The outline of this paper is as follows. First, we give a brief overview of atomic structure (Sect. 2). We then discuss a few basic processes that play an important role for the thermal plasmas considered here (Sect. 3). For the proper calculation of an X-ray spectrum, three different steps should be considered:

1. the determination of the ionisation balance (Sect. 4),
2. the determination of the emitted spectrum (Sect. 5),
3. possible absorption of the photons on their way towards Earth (Sect. 6).

We then briefly discuss issues like the Galactic foreground emission (Sect. 7), plasma cooling (Sect. 8), the role of non-thermal electrons (Sect. 9), and conclude with a section on plasma modelling (Sect. 10).

2 A Short Introduction to Atomic Structure

2.1 The Bohr Atom

The electrons in an atom have orbits with discrete energy levels and quantum numbers. The principal quantum number n corresponds to the energy I_n of the orbit (in the classical Bohr model for the hydrogen atom the energy $I_n = E_{\rm H} n^{-2}$ with $E_{\rm H} = 13.6$ eV the Rydberg energy), and it takes discrete values $n = 1, 2, 3, \ldots$. An atomic shell consists of all electrons with the same value of n.

The second quantum number ℓ corresponds to the angular momentum of the electron, and takes discrete values $\ell < n$. Orbits with $\ell = 0, 1, 2, 3$ are designated as s, p, d, and f orbits. A subshell consists of all electrons with the same value of n and ℓ; they are usually designated as 1s, 2s, 2p, etc.

The spin quantum number s of an electron can take values $s = \pm 1/2$, and the combined total angular momentum j has a quantum number with values between $\ell - 1/2$ (for $\ell > 0$) and $\ell + 1/2$. Subshells are subdivided according to their j-value and are designated as $n\ell_j$. Example: $n = 2$, $\ell = 1$, $j = 3/2$ is designated as $2{\rm p}_{3/2}$.

There is also another notation that is commonly used in X-ray spectroscopy. Shells with $n = 1, 2, 3, 4, 5, 6$ and 7 are indicated with K, L, M, N, O, P, Q. A further subdivision is

made starting from low values of ℓ up to higher values of ℓ and from low values of j up to higher values of j:

1s	2s	$2p_{1/2}$	$2p_{3/2}$	3s	$3p_{1/2}$	$3p_{3/2}$	$3d_{3/2}$	$3d_{5/2}$	4s	etc.
K	L_I	L_{II}	L_{III}	M_I	M_{II}	M_{III}	M_{IV}	M_V	N_I	
2	2	2	4	2	2	4	4	6	2	

The third row in this table indicates the maximum number of electrons that can be contained in each subshell. This so-called statistical weight is simply $2j + 1$.

Atoms or ions with multiple electrons in most cases have their shells filled starting from low to high n and ℓ. For example, neutral oxygen has 8 electrons, and the shells are filled like $1s^2 2s^2 2p^4$, where the superscripts denote the number of electrons in each shell. Ions or atoms with completely filled subshells (combining all allowed j-values) are particularly stable. Examples are the noble gases: neutral helium, neon and argon, and more general all ions with 2, 10 or 18 electrons. In chemistry, it is common practise to designate ions by the number of electrons that they have lost, like O^{2+} for doubly ionised oxygen. In astronomical spectroscopic practise, more often one starts to count with the neutral atom, such that O^{2+} is designated as O III. As the atomic structure and the possible transitions of an ion depend primarily on the number of electrons it has, there are all kinds of scaling relationships along so-called iso-electronic sequences. All ions on an iso-electronic sequence have the same number of electrons; they differ only by the nuclear charge Z. Along such a sequence, energies, transitions or rates are often continuous functions of Z. A well known example is the famous 1s−2p triplet of lines in the helium iso-electronic sequence (2-electron systems), e.g. C V, N VI and O VII.

2.2 Level Notation in Multi-Electron Systems

For ions or atoms with more than one electron, the atomic structure is determined by the combined quantum numbers of the electrons. These quantum numbers have to be added according to the rules of quantum mechanics. We will not go into detail here, but refer to textbooks such as Herzberg (1944). The determination of the allowed quantum states and the transitions between these states can be rather complicated. Important here is to know that for each electron configuration (for example 1s 2p) there is a number of allowed terms or multiplets, designated as ^{2S+1}L, with S the combined electron spin (derived from the individual s values of the electrons), and L represents the combined angular momentum (from the individual ℓ values). For L one usually substitutes the alphabetic designations similar to those for single electrons, namely S for $L = 0$, P for $L = 1$, etc. The quantity $2S + 1$ represents the multiplicity of the term, and gives the number of distinct energy levels of the term. The energy levels of each term can be distinguished by J, the combined total angular momentum quantum number j of the electrons. Terms with $2S + 1$ equals 1, 2 or 3 are designated as singlets, doublets and triplets, etc.

For example, a configuration with two equivalent p electrons (e.g., $3p^2$), has three allowed terms, namely ^{1}S, ^{1}D and ^{3}P. The triplet term ^{3}P has $2S + 1 = 3$ hence $S = 1$ and $L = 1$ (corresponding to P), and the energy levels of this triplet are designated as $^3P_0$, $^3P_1$ and $^3P_2$ according to their J values 0, 1 and 2.

2.3 Binding Energies

The binding energy I of K-shell electrons in neutral atoms increases approximately as $I \sim Z^2$ with Z being the nuclear charge (see Table 1 and Fig. 1). Also for other shells the energy increases strongly with increasing nuclear charge Z.

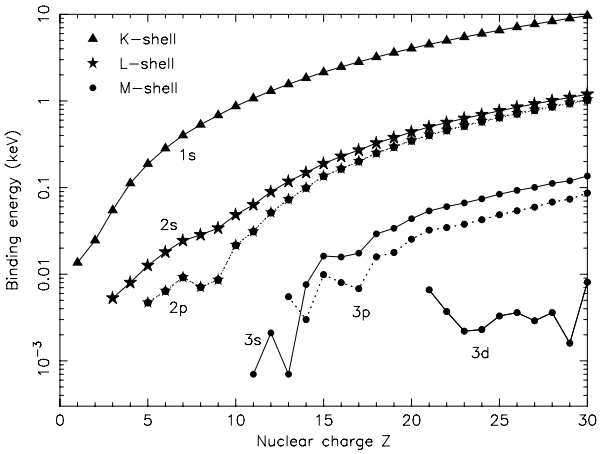


Fig. 1 Energy levels of atomic subshells for neutral atoms

Table 1 Binding energies E, corresponding wavelengths λ and photoionisation cross sections σ at the edges of neutral atoms (nuclear charge Z) for abundant elements. Only edges in the X-ray band are given. Note that 1 Mbarn $= 10^{-22}$ m^2

	Z	E (eV)	λ (Å)	σ (Mbarn)		Z	E (eV)	λ (Å)	σ (Mbarn)
1s shell:					2s shell:				
H	1	13.6	911.8	6.29	O	8	16.6	747.3	1.37
He	2	24.6	504.3	7.58	Si	14	154	80.51	0.48
C	6	288	43.05	0.96	S	16	232	53.44	0.37
N	7	403	30.77	0.67	Ar	18	327	37.97	0.29
O	8	538	23.05	0.50	Ca	20	441	28.11	0.25
Ne	10	870	14.25	0.29	Fe	26	851	14.57	0.15
Mg	12	1308	9.48	0.20	Ni	28	1015	12.22	0.13
Si	14	1844	6.72	0.14	$2p_{1/2}$ shell:				
S	16	2476	5.01	0.096	S	16	169	73.26	1.61
Ar	18	3206	3.87	0.070	Ar	18	251	49.46	1.45
Ca	20	4041	3.07	0.060	Ca	20	353	35.16	0.86
Fe	26	7117	1.74	0.034	Fe	26	726	17.08	0.42
Ni	28	8338	1.49	0.029	Ni	28	877	14.13	0.35
					$2p_{3/2}$ shell:				
					S	16	168	73.80	3.24
					Ar	18	249	49.87	2.97
					Ca	20	349	35.53	1.74
					Fe	26	713	17.39	0.86
					Ni	28	860	14.42	0.73

For ions of a given element the ionisation energies decrease with decreasing ionisation stage: for lowly ionised ions, a part of the Coulomb force exerted on an electron by the positively charged nucleus is compensated by the other electrons in the ion, thereby allowing a wider orbit with less energy. An example is given in Table 2.

Table 2 Binding energies E, corresponding wavelengths λ and photoionisation cross sections σ at the K-edges of oxygen ions

Ion	E (eV)	λ (Å)	σ (Mbarn) (10^{-22} m^2)
O I	544	22.77	0.50
O II	565	21.94	0.45
O III	592	20.94	0.41
O IV	618	20.06	0.38
O V	645	19.22	0.35
O VI	671	18.48	0.32
O VII	739	16.77	0.24
O VIII	871	14.23	0.10

Table 3 Proto-solar abundances for the 15 most common chemical elements. Abundances A are given with respect to hydrogen. Data from Lodders (2003)

Element	Abundance	Element	Abundance	Element	Abundance
H	$\equiv 1$	Ne	89.1×10^{-6}	S	18.2×10^{-6}
He	0.0954	Na	2.34×10^{-6}	Ar	4.17×10^{-6}
C	288×10^{-6}	Mg	41.7×10^{-6}	Ca	2.57×10^{-6}
N	79.4×10^{-6}	Al	3.47×10^{-6}	Fe	34.7×10^{-6}
O	575×10^{-6}	Si	40.7×10^{-6}	Ni	1.95×10^{-6}

2.4 Abundances

With high spectral resolution and sensitivity, optical spectra of stars sometimes show spectral features from almost all elements of the Periodic Table, but in practise only a few of the most abundant elements show up in X-ray spectra of cosmic plasmas. In several situations the abundances of the chemical elements in an X-ray source are similar to (but not necessarily equal to) the abundances for the Sun or a constant fraction of that. There have been many significant changes over the last several years in the adopted set of standard cosmic abundances. A still often used set of abundances is that of Anders and Grevesse (1989), but a more recent one is the set of proto-solar abundances of Lodders (2003), that we list in Table 3 for a few of the key elements.

In general, for absorption studies the strength of the lines is mainly determined by atomic parameters that do not vary much along an iso-electronic sequence, and the abundance of the element. Therefore, in the X-ray band the oxygen lines are the strongest absorption lines, as oxygen is the most abundant metal. The emissivity of ions in the X-ray band often increases with a strong power of the nuclear charge. For that reason, in many X-ray plasmas the strongest iron lines are often of similar strength to the strongest oxygen lines, despite the fact that the cosmic abundance of iron is only 6% of the cosmic oxygen abundance.

3 Basic Processes

3.1 Excitation Processes

A bound electron in an ion can be brought into a higher, excited energy level through a collision with a free electron or by absorption of a photon. The latter will be discussed in more detail in Sect. 6. Here we focus upon excitation by electrons.

The cross section Q_{ij} for excitation from level i to level j for this process can be conveniently parametrised by

$$Q_{ij}(U) = \frac{\pi a_0^2}{w_i} \frac{E_\mathrm{H}}{E_{ij}} \frac{\Omega(U)}{U}, \tag{1}$$

where $U = E_{ij}/E$ with E_{ij} the excitation energy from level i to j, E the energy of the exciting electron, E_H the Rydberg energy (13.6 eV), a_0 the Bohr radius and w_i the statistical weight of the lower level i. The dimensionless quantity $\Omega(U)$ is the so-called collision strength. For a given transition on an iso-electronic sequence, $\Omega(U)$ is not a strong function of the atomic number Z, or may be even almost independent of Z.

Mewe (1972) introduced a convenient formula that can be used to describe most collision strengths, written here as follows:

$$\Omega(U) = A + \frac{B}{U} + \frac{C}{U^2} + \frac{2D}{U^3} + F \ln U, \tag{2}$$

where A, B, C, D and F are parameters that differ for each transition. The expression can be integrated analytically over a Maxwellian electron distribution, and the result can be expressed in terms of exponential integrals. We only mention here that the total excitation rate S_{ij} (in units of $\mathrm{m}^{-3}\,\mathrm{s}^{-1}$) is given by

$$S_{ij} = 8.62 \times 10^{-12} \frac{\bar{\Omega}(y) T^{-1/2} \mathrm{e}^{-y}}{w_i}, \tag{3}$$

with $y \equiv E_{ij}/\mathrm{k}T$ and $\bar{\Omega}(y)$ is the Maxwellian-averaged collision strength. For low temperatures, $y \gg 1$ and $\bar{\Omega}(y) = A + B + C + 2D$, leading to $S_{ij} \sim T^{-1/2}\mathrm{e}^{-y}$. The excitation rate drops exponentially due to the lack of electrons with sufficient energy. In the other limit of high temperature, $y \ll 1$ and $\bar{\Omega}(y) = -F \ln y$ and hence $S_{ij} \sim T^{-1/2} \ln y$.

Not all transitions have this asymptotic behaviour, however. For instance, so-called forbidden transitions have $F = 0$ and hence have much lower excitation rates at high energy. So-called spin-forbidden transitions even have $A = B = F = 0$.

In most cases, the excited state is stable and the ion will decay back to the ground level by means of a radiative transition, either directly or through one or more steps via intermediate energy levels. Only in cases of high density or high radiation fields, collisional excitation or further radiative excitation to a higher level may become important, but for typical cluster and ISM conditions these processes are not important in most cases. Therefore, the excitation rate immediately gives the total emission line power.

3.2 Ionisation Processes

3.2.1 Collisional Ionisation

Collisional ionisation occurs when during the interaction of a free electron with an atom or ion the free electron transfers a part of its energy to one of the bound electrons, which is

then able to escape from the ion. A necessary condition is that the kinetic energy E of the free electron must be larger than the binding energy I of the atomic shell from which the bound electron escapes. A formula that gives a correct order of magnitude estimate of the cross section σ of this process and that has the proper asymptotic behaviour (first calculated by Bethe and Born) is the formula of Lotz (1968):

$$\sigma = \frac{a n_s \ln(E/I)}{E I}, \tag{4}$$

where n_s is the number of electrons in the shell and the normalisation $a = 4.5 \times 10^{-24}$ m^2 keV^2. This equation shows that high-energy electrons have less ionising power than low-energy electrons. Also, the cross section at the threshold $E = I$ is zero.

The above cross section has to be averaged over the electron distribution (Maxwellian for a thermal plasma). For simple formulae for the cross section such as (4) the integration can be done analytically and the result can be expressed in terms of exponential integrals. We give here only the asymptotic results for $C_{\rm DI}$, the total number of direct ionisations per unit volume per unit time:

$$\mathrm{k}T \ll I: \quad C_{\rm DI} \simeq \left(\frac{2\sqrt{2} a n_s}{\sqrt{\pi m_{\rm e}}}\right) \frac{n_{\rm e} n_{\rm i} \sqrt{\mathrm{k}T}\, \mathrm{e}^{-I/\mathrm{k}T}}{I^2}, \tag{5}$$

and

$$\mathrm{k}T \gg I: \quad C_{\rm DI} \simeq \left(\frac{2\sqrt{2} a n_s}{\sqrt{\pi m_{\rm e}}}\right) \frac{n_{\rm e} n_{\rm i} \ln(\mathrm{k}T/I)}{I \sqrt{\mathrm{k}T}}. \tag{6}$$

For low temperatures, the ionisation rate therefore goes exponentially to zero. This can be understood simply, because for low temperatures only the electrons from the exponential tail of the Maxwell distribution have sufficient energy to ionise the atom or ion. For higher temperatures the ionisation rate also approaches zero, but this time because the cross section at high energies is small.

For each ion the direct ionisation rate per atomic shell can now be determined. The total rate follows immediately by adding the contributions from the different shells. Usually only the outermost two or three shells are important. That is because of the scaling with I^{-2} and I^{-1} in (5) and (6), respectively.

3.2.2 Photoionisation

This process is very similar to collisional ionisation. The difference is that in the case of photoionisation a photon instead of an electron is causing the ionisation. Further, the effective cross section differs from that for collisional ionisation. As an example Fig. 2 shows the cross section for neutral iron and Na-like iron. For the more highly ionised iron the so-called "edges" (corresponding to the ionisation potentials I) occur at higher energies than for neutral iron. Contrary to the case of collisional ionisation, the cross section at threshold for photoionisation is not zero. The effective cross section just above the edges sometimes changes rapidly (the so-called Cooper minima and maxima).

Contrary to collisional ionisation, all the inner shells now have the largest cross section. For the K-shell one can approximate for $E > I$

$$\sigma(E) \simeq \sigma_0 (I/E)^3. \tag{7}$$

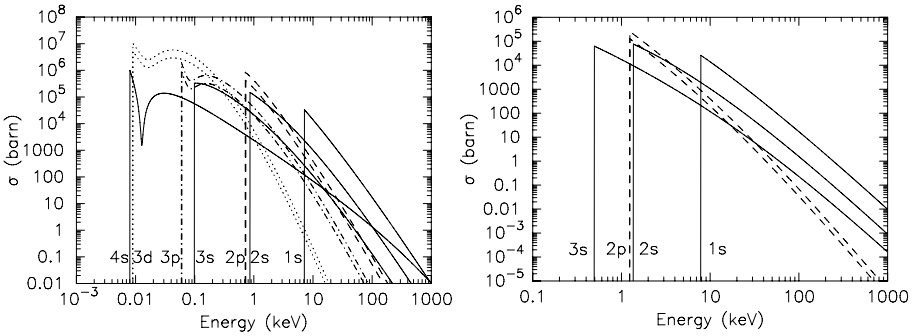


Fig. 2 Photoionisation cross section in barn (10^{-28} m^{-2}) for Fe I (*left*) and Fe XVI (*right*). The p and d states (*dashed* and *dotted*) have two lines each because of splitting into two sublevels with different j (see Sect. 2.1)

For a given ionising spectrum $F(E)$ (photons per unit volume per unit energy) the total number of photoionisations follows as

$$C_{\rm PI} = c \int_0^\infty n_{\rm i} \sigma(E) F(E) {\rm d}E. \tag{8}$$

For hydrogenlike ions one can write:

$$\sigma_{\rm PI} = \frac{64\pi n g(E,n) \alpha a_0^2}{3\sqrt{3} Z^2} \left(\frac{I}{E}\right)^3, \tag{9}$$

where n is the principal quantum number, α the fine structure constant and a_0 the Bohr radius. The Gaunt factor $g(E,n)$ is of order unity and varies only slowly as a function of E. It has been calculated and tabulated by Karzas and Latter (1961). The above equation is also applicable to excited states of the atom, and is a good approximation for all excited atoms or ions where n is larger than the corresponding value for the valence electron.

3.2.3 Compton Ionisation

Scattering of a photon on an electron generally leads to energy transfer from one of the particles to the other. In most cases only scattering on free electrons is considered. But Compton scattering also can occur on bound electrons. If the energy transfer from the photon to the electron is large enough, the ionisation potential can be overcome leading to ionisation. This is the Compton ionisation process.

In the Thomson limit the differential cross section for Compton scattering is given by

$$\frac{{\rm d}\sigma}{{\rm d}\Omega} = \frac{3\sigma_{\rm T}}{16\pi}(1+\cos^2\theta), \tag{10}$$

with θ the scattering angle and $\sigma_{\rm T}$ the Thomson cross section (6.65×10^{-29} m^{-2}). The energy transfer ΔE during the scattering is given by (E is the photon energy):

$$\Delta E = \frac{E^2(1-\cos\theta)}{m_{\rm e}c^2 + E(1-\cos\theta)}. \tag{11}$$

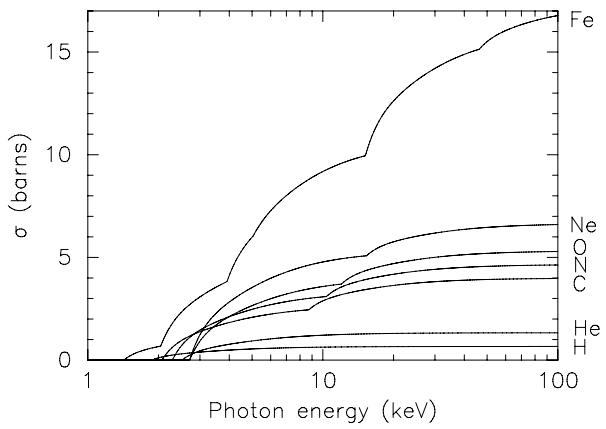


Fig. 3 Compton ionisation cross section for neutral atoms of H, He, C, N, O, Ne and Fe

Only those scatterings where $\Delta E > I$ contribute to the ionisation. This defines a critical angle θ_c, given by:

$$\cos\theta_c = 1 - \frac{m_e c^2 I}{E^2 - IE}. \tag{12}$$

For $E \gg I$ we have $\sigma(E) \to \sigma_T$ (all scatterings lead in that case to ionisation) and further for $\theta_c \to \pi$ we have $\sigma(E) \to 0$. Because for most ions $I \ll m_e c^2$, this last condition occurs for $E \simeq \sqrt{I m_e c^2/2} \gg I$. See Fig. 3 for an example of some cross sections. In general, Compton ionisation is important if the ionising spectrum contains a significant hard X-ray contribution, for which the Compton cross section is larger than the steeply falling photoionisation cross section.

3.2.4 Autoionisation and Fluorescence

As we showed above, interaction of a photon or free electron with an atom or ion may lead to ionisation. In particular when an electron from one of the inner shells is removed, the resulting ion has a "vacancy" in its atomic structure and is unstable. Two different processes may occur to stabilise the ion again.

The first process is fluorescence. Here one of the electrons from the outer shells makes a radiative transition in order to occupy the vacancy. The emitted photon has an energy corresponding to the energy difference between the initial and final discrete states.

The other possibility to fill the gap is auto-ionisation through the Auger process. In this case, also one of the electrons from the outer shells fills the vacancy in the lower level. The released energy is not emitted as a photon, however, but transferred to another electron from the outer shells that is therefore able to escape from the ion. As a result, the initial ionisation may lead to a double ionisation. If the final electron configuration of the ion still has holes, more auto-ionisations or fluorescence may follow until the ion has stabilised.

In Fig. 4 the fluorescence yield ω (probability that a vacancy will be filled by a radiative transition) is shown for all elements. In general, the fluorescence yield increases strongly with increasing nuclear charge Z, and is higher for the innermost atomic shells. As a typical example, for Fe I a K-shell vacancy has $\omega = 0.34$, while an L_I-shell vacancy has $\omega = 0.001$. For O I these numbers are 0.009 and 0, respectively.

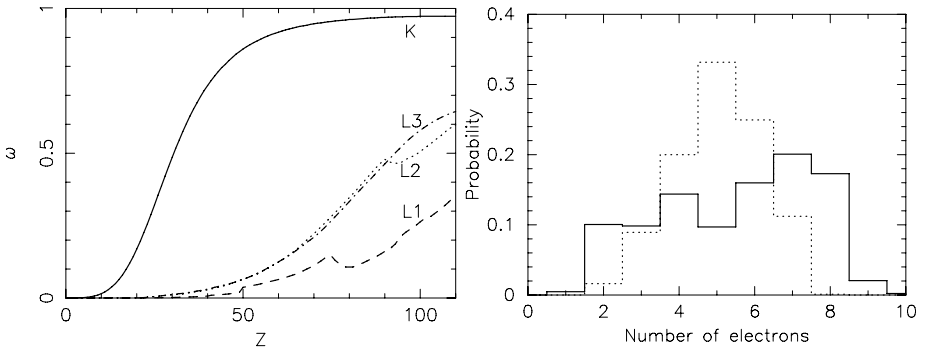


Fig. 4 *Left panel*: Fluorescence yield ω as a function of atomic number Z for the K and L shells. *Right panel*: Distribution of number of electrons liberated after the initial removal of an electron from the K-shell (*solid line*) or L_I shell (*dotted line*), including the original photo-electron, for Fe I; after Kaastra and Mewe (1993)

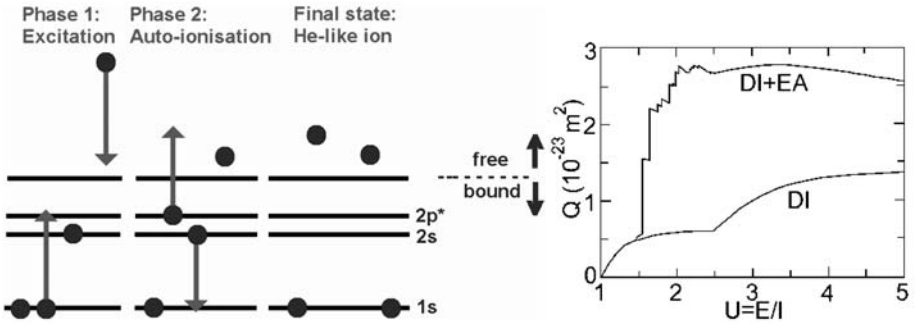


Fig. 5 *Left panel*: the Excitation-Autoionisation process for a Li-like ion. *Right panel*: collisional ionisation cross section for Fe XVI. The contribution from direct collisional ionisation (DI) and excitation-autoionisation (EA) are indicated. Adapted from Arnaud and Rothenflug (1985)

3.2.5 Excitation-Autoionisation

In Sect. 3.2.1 we showed how the collision of a free electron with an ion can lead to ionisation. In principle, if the free electron has insufficient energy ($E < I$), there will be no ionisation. However, even in that case it is sometimes still possible to ionise, but in a more complicated way. The process is as follows. The collision can bring one of the electrons in the outermost shells in a higher quantum level (excitation). In most cases, the ion will return to its ground level by a radiative transition. But in some cases the excited state is unstable, and a radiationless Auger transition can occur (see Sect. 3.2.4). The vacancy that is left behind by the excited electron is being filled by another electron from one of the outermost shells, while the excited electron is able to leave the ion (or a similar process with the role of both electrons reversed).

Because of energy conservation, this process only occurs if the excited electron comes from one of the inner shells (the ionisation potential barrier must be taken anyhow). The process is in particular important for Li-like and Na-like ions, and for several other individual atoms and ions. As an example we treat here Li-like ions (see Fig. 5, left panel). In that case the most important contribution comes from a 1s−2p excitation.

3.3 Recombination Processes

3.3.1 Radiative Recombination

Radiative recombination is the reverse process of photoionisation. A free electron is captured by an ion while emitting a photon. The released radiation is the so-called free-bound continuum emission. It is relatively easy to show that there is a simple relation between the photoionisation cross section $\sigma_{\rm bf}(E)$ and the recombination cross section $\sigma_{\rm fb}$, namely the Milne-relation:

$$\sigma_{\rm fb}(v) = \frac{E^2 g_n \sigma_{\rm bf}(E)}{m_{\rm e} c^2 m_{\rm e} v^2} \tag{13}$$

where g_n is the statistical weight of the quantum level into which the electron is captured (for an empty shell this is $g_n = 2n^2$). By averaging over a Maxwell distribution one gets the recombination-coefficient to level n:

$$R_n = n_{\rm i} n_{\rm e} \int_0^\infty v f(v) \sigma_{\rm fb}(v) {\rm d}v. \tag{14}$$

Of course there is energy conservation, so $E = \frac{1}{2} m_{\rm e} v^2 + I$.

It can be shown that for the photoionisation cross section (9) and for $g = 1$, constant and $g_n = 2n^2$:

$$R_n = \frac{128\sqrt{2\pi} n^3 \alpha a_0^2 I^3 {\rm e}^{I/{\rm k}T} n_{\rm i} n_{\rm e}}{3\sqrt{3 m_{\rm e} {\rm k}T}\, Z^2 {\rm k}T m_{\rm e} {\rm c}^3} E_1(I/{\rm k}T). \tag{15}$$

With the asymptotic relations for the exponential integrals it can be shown that

$${\rm k}T \ll I: \quad R_n \sim T^{-\frac{1}{2}}, \tag{16}$$

$${\rm k}T \gg I: \quad R_n \sim \ln(I/{\rm k}T) T^{-3/2}. \tag{17}$$

Therefore for $T \to 0$ the recombination coefficient approaches infinity: a cool plasma is hard to ionise. For $T \to \infty$ the recombination coefficient goes to zero, because of the Milne relation ($v \to \infty$) and because of the sharp decrease of the photoionisation cross section for high energies.

As a rough approximation we can use further that $I \sim (Z/n)^2$. Substituting this we find that for ${\rm k}T \ll I$ (recombining plasmas) $R_n \sim n^{-1}$, while for ${\rm k}T \gg I$ (ionising plasmas) $R_n \sim n^{-3}$. In recombining plasmas in particular many higher excited levels will be populated by the recombination, leading to significantly stronger line emission. On the other hand, in ionising plasmas (such as supernova remnants) recombination mainly occurs to the lowest levels. Note that for recombination to the ground level the approximation (15) cannot be used (the hydrogen limit), but instead one should use the exact photoionisation cross section of the valence electron. By adding over all values of n and applying an approximation Seaton (1959) found for the total radiative recombination rate $\alpha_{\rm RR}$ (in units of ${\rm m}^{-3}\ {\rm s}^{-1}$):

$$\alpha_{\rm RR} \equiv \sum_n R_n = 5.197 \times 10^{-20}\ Z \lambda^{1/2} \{0.4288 + 0.5 \ln \lambda + 0.469 \lambda^{-1/3}\} \tag{18}$$

with $\lambda \equiv E_{\rm H} Z^2 / {\rm k}T$ and $E_{\rm H}$ the Rydberg energy (13.6 eV). Note that this equation only holds for hydrogen-like ions. For other ions usually an analytical fit to numerical calculations is

used:

$$\alpha_{\rm RR} \sim T^{-\eta} \tag{19}$$

where the approximation is strictly speaking only valid for T near the equilibrium concentration. The approximations (16) and (17) raise suspicion that for $T \to 0$ or $T \to \infty$ (19) could be a poor choice.

The captured electron does not always reach the ground level immediately. We have seen before that in particular for cool plasmas ($kT \ll I$) the higher excited levels are frequently populated. In order to get to the ground state, one or more radiative transitions are required. Apart from cascade corrections from and to higher levels the recombination line radiation is essentially given by (15). A comparison of recombination with excitation tells that in particular for low temperatures (compared to the line energy) recombination radiation dominates, and for high temperatures excitation radiation dominates. This is also one of the main differences between photoionised and collisionally ionised plasmas, as photoionised plasmas in general have a low temperature compared to the typical ionisation potentials.

3.3.2 Dielectronic Recombination

This process is more or less the inverse of excitation-autoionisation. Now a free electron interacts with an ion, by which it is caught (quantum level $n''\ell''$) but at the same time it excites an electron from $(n\ell) \to (n'\ell')$. The doubly excited state is in general not stable, and the ion will return to its original state by auto-ionisation. However there is also a possibility that one of the excited electrons (usually the electron that originally belonged to the ion) falls back by a radiative transition to the ground level, creating therefore a stable, albeit excited state $(n''\ell'')$ of the ion. In particular excitations with $\ell' = \ell + 1$ contribute much to this process. In order to calculate this process, one should take account of many combinations $(n'\ell')(n''\ell'')$.

The final transition probability is often approximated by

$$\alpha_{\rm DR} = \frac{A}{T^{3/2}} {\rm e}^{-T_0/T} (1 + B{\rm e}^{-T_1/T}) \tag{20}$$

where A, B, T_0 and T_1 are adjustable parameters. Note that for $T \to \infty$ the asymptotic behaviour is identical to the case of radiative recombination. For $T \to 0$ however, dielectronic recombination can be neglected; this is because the free electron has insufficient energy to excite a bound electron. Dielectronic recombination is a dominant process in the Solar corona, and also in other situations it is often very important.

Dielectronic recombination produces more than one line photon. Consider for example the dielectronic recombination of a He-like ion into a Li-like ion:

$$\mathrm{e} + 1\mathrm{s}^2 \to 1\mathrm{s}2\mathrm{p}3\mathrm{s} \to 1\mathrm{s}^2 3\mathrm{s} + \mathrm{h}\nu_1 \to 1\mathrm{s}^2 2\mathrm{p} + \mathrm{h}\nu_2 \to 1\mathrm{s}^2 2\mathrm{s} + \mathrm{h}\nu_3 \tag{21}$$

The first arrow corresponds to the electron capture, the second arrow to the stabilising radiative transition 2p→1s and the third arrow to the radiative transition 3s→2p of the captured electron. This last transition would have also occurred if the free electron was caught directly into the 3s shell by normal radiative recombination. Finally, the electron has to decay further to the ground level and this can go through the normal transitions in a Li-like ion (fourth arrow). This single recombination thus produces three line photons.

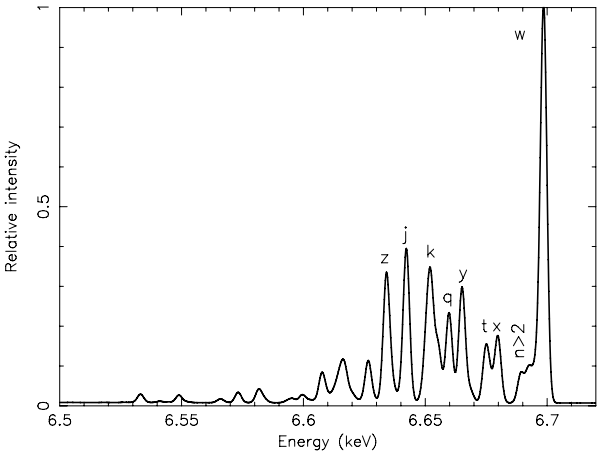


Fig. 6 Spectrum of a plasma in collisional ionisation equilibrium with $kT = 2$ keV, near the Fe-K complex. Lines are labelled using the most common designations in this field. The Fe XXV "triplet" consists of the resonance line (w), intercombination line (actually split into x and y) and the forbidden line (z). All other lines are satellite lines. The labelled satellites are lines from Fe XXIV, most of the lines with energy below the forbidden (z) line are from Fe XXIII. The relative intensity of these satellites is a strong indicator for the physical conditions in the source

Because of the presence of the extra electron in the higher orbit, the energy $h\nu_1$ of the 2p→1s transition is slightly different from the energy in a normal He-like ion. The stabilising transition is therefore also called a satellite line. Because there are many different possibilities for the orbit of the captured electron, one usually finds a forest of such satellite lines surrounding the normal 2p→1s excitation line in the He-like ion (or analogously for other iso-electronic sequences). Fig. 6 gives an example of these satellite lines.

3.4 Charge Transfer Processes

In most cases ionisation or recombination in collisionally ionised plasmas is caused by interactions of an ion with a free electron. At low temperatures (typically below 10^5 K) also charge transfer reactions become important. During the interaction of two ions, an electron may be transferred from one ion to the other; it is usually captured in an excited state, and the excited ion is stabilised by one or more radiative transitions. As hydrogen and helium outnumber by at least a factor of 100 any other element (see Table 3), in practise only interactions between those elements and heavier nuclei are important. Reactions with H I and He I lead to recombination of the heavier ion, and reactions with H II and He II to ionisation.

The electron captured during charge transfer recombination of an oxygen ion (for instance O VII, O VIII) is usually captured in an intermediate quantum state (principal quantum number $n = 4$–6). This leads to enhanced line emission from that level as compared to the emission from other principal levels, and this signature can be used to show the presence of charge transfer reactions. Another signature – actually a signature for all recombining plasmas – is of course the enhancement of the forbidden line relative to the resonance line in the O VII triplet (Sect. 5.2.2).

An important example is the charge transfer of highly charged ions from the Solar wind with the neutral or weakly ionised Geocorona. Whenever the Sun is more active, this process may produce enhanced thermal soft X-ray emission in addition to the steady foreground emission from our own Galaxy. See Bhardwaj et al. (2006) for a review of X-rays from the Solar System. Care should be taken not to confuse this temporary enhanced Geocoronal emission with soft excess emission in a background astrophysical source.

4 Ionisation Balance

In order to calculate the X-ray emission or absorption from a plasma, apart from the physical conditions also the ion concentrations must be known. These ion concentrations can be determined by solving the equations for ionisation balance (or in more complicated cases by solving the time-dependent equations). A basic ingredient in these equations are the ionisation and recombination rates, that we discussed in the previous section. Here we consider three of the most important cases: collisional ionisation equilibrium, non-equilibrium ionisation and photoionisation equilibrium.

4.1 Collisional Ionisation Equilibrium (CIE)

The simplest case is a plasma in collisional ionisation equilibrium (CIE). In this case one assumes that the plasma is optically thin for its own radiation, and that there is no external radiation field that affects the ionisation balance.

Photo-ionisation and Compton ionisation therefore can be neglected in the case of CIE. This means that in general each ionisation leads to one additional free electron, because the direct ionisation and excitation-autoionisation processes are most efficient for the outermost atomic shells. The relevant ionisation processes are collisional ionisation and excitation-autoionisation, and the recombination processes are radiative recombination and dielectronic recombination. Apart from these processes, at low temperatures also charge transfer ionisation and recombination are important.

We define R_z as the total recombination rate of an ion with charge z to charge $z-1$, and I_z as the total ionisation rate for charge z to $z+1$. Ionisation equilibrium then implies that the net change of ion concentrations n_z should be zero:

$$z > 0: \quad n_{z+1}R_{z+1} - n_z R_z + n_{z-1}I_{z-1} - n_z I_z = 0 \tag{22}$$

and in particular for $z = 0$ one has

$$n_1 R_1 = n_0 I_0 \tag{23}$$

(a neutral atom cannot recombine further and it cannot be created by ionisation). Next an arbitrary value for n_0 is chosen, and (23) is solved:

$$n_1 = n_0(I_0/R_1). \tag{24}$$

This is substituted into (22) which now can be solved. Using induction, it follows that

$$n_{z+1} = n_z(I_z/R_{z+1}). \tag{25}$$

Finally everything is normalised by demanding that

$$\sum_{z=0}^{Z} n_z = n_{\text{element}} \tag{26}$$

where n_{element} is the total density of the element, determined by the total plasma density and the chemical abundances.

Examples of plasmas in CIE are the Solar corona, coronae of stars, the hot intracluster medium, the diffuse Galactic ridge component. Figure 7 shows the ion fractions as a function of temperature for two important elements.

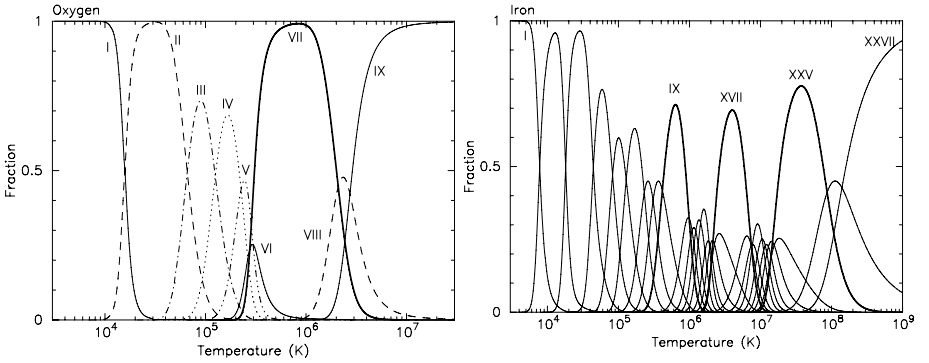


Fig. 7 Ion concentration of oxygen ions (*left panel*) and iron ions (*right panel*) as a function of temperature in a plasma in Collisional Ionisation Equilibrium (CIE). Ions with completely filled shells are indicated with *thick lines*: the He-like ions O VII and Fe XXV, the Ne-like Fe XVII and the Ar-like Fe IX; note that these ions are more prominent than their neighbours

4.2 Non-Equilibrium Ionisation (NEI)

The second case that we discuss is non-equilibrium ionisation (NEI). This situation occurs when the physical conditions of the source, like the temperature, suddenly change. A shock, for example, can lead to an almost instantaneous rise in temperature. However, it takes a finite time for the plasma to respond to the temperature change, as ionisation balance must be recovered by collisions. Similar to the CIE case we assume that photoionisation can be neglected. For each element with nuclear charge Z we write:

$$\frac{1}{n_\mathrm{e}(t)}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{n}(Z,t) = \mathbf{A}(Z,T(t))\mathbf{n}(Z,t) \tag{27}$$

where $\mathbf{n}$ is a vector of length $Z+1$ that contains the ion concentrations, and which is normalised according to (26). The transition matrix $\mathbf{A}$ is a $(Z+1)\times(Z+1)$ matrix given by

$$\mathbf{A} = \begin{pmatrix} -I_0 & R_1 & 0 & 0 & \dots & \\ I_0 & -(I_1+R_1) & R_2 & 0 & & \\ 0 & I_1 & \dots & \dots & & \\ & \vdots & \ddots & \vdots & \dots & \\ & & \dots & R_{Z-1} & 0 & \\ & \dots & 0 & I_{Z-2} & -(I_{Z-1}+R_{Z-1}) & R_Z \\ & & \dots & 0 & I_{Z-1} & -R_Z \end{pmatrix}.$$

We can write the equation in this form because both ionisations and recombinations are caused by collisions of electrons with ions. Therefore we have the uniform scaling with n_e. In general, the set of equations (27) must be solved numerically. The time evolution of the plasma can be described in general well by the parameter

$$U = \int n_\mathrm{e}\mathrm{d}t. \tag{28}$$

The integral should be done over a co-moving mass element. Typically, for most ions equilibrium is reached for $U \sim 10^{18}$ m^{-3} s. We should mention here, however, that the final outcome also depends on the temperature history $T(t)$ of the mass element, but in most cases the situation is simplified to $T(t) =$ constant.

4.3 Photoionisation Equilibrium (PIE)

The third case that we treat are the photoionised plasmas. Usually one assumes equilibrium (PIE), but there are of course also extensions to non-equilibrium photo-ionised situations. Apart from the same ionisation and recombination processes that play a role for plasmas in NEI and CIE, also photoionisation and Compton ionisation are relevant. Because of the multiple ionisations caused by Auger processes, the equation for the ionisation balance is not as simple as (22), because now one needs to couple more ions. Moreover, not all rates scale with the product of electron and ion density, but the balance equations also contain terms proportional to the product of ion density times photon density. In addition, apart from the equation for the ionisation balance, one needs to solve simultaneously an energy balance equation for the electrons. In this energy equation not only terms corresponding to ionisation and recombination processes play a role, but also several radiation processes (Bremsstrahlung, line radiation) or Compton scattering. The equilibrium temperature must be determined in an iterative way. A classical paper describing such photoionised plasmas is Kallman and McCray (1982).

5 Emission Processes

5.1 Continuum Emission Processes

In any plasma there are three important continuum emission processes, that we briefly mention here: Bremsstrahlung, free-bound emission and two-photon emission.

5.1.1 Bremsstrahlung

Bremsstrahlung is caused by a collision between a free electron and an ion. The emissivity $\epsilon_{\rm ff}$ (photons m^{-3} s^{-1} J^{-1}) can be written as:

$$\epsilon_{\rm ff} = \frac{2\sqrt{2}\alpha\sigma_{\rm T} c n_{\rm e} n_{\rm i} Z_{\rm eff}^2}{\sqrt{3\pi}E}\left(\frac{m_{\rm e}c^2}{{\rm k}T}\right)^{\frac{1}{2}} g_{\rm ff} e^{-E/{\rm k}T}\,, \tag{29}$$

where α is the fine structure constant, $\sigma_{\rm T}$ the Thomson cross section, $n_{\rm e}$ and $n_{\rm i}$ the electron and ion density, and E the energy of the emitted photon. The factor $g_{\rm ff}$ is the so-called Gaunt factor and is a dimensionless quantity of order unity. Further, $Z_{\rm eff}$ is the effective charge of the ion, defined as

$$Z_{\rm eff} = \left(\frac{n_r^2 I_r}{E_{\rm H}}\right)^{\frac{1}{2}} \tag{30}$$

where $E_{\rm H}$ is the ionisation energy of hydrogen (13.6 eV), I_r the ionisation potential of the ion after a recombination, and n_r the corresponding principal quantum number.

It is also possible to write (29) as $\epsilon_{\rm ff} = P_{\rm ff} n_{\rm e} n_{\rm i}$ with

$$P_{\rm ff} = \frac{3.031 \times 10^{-21} Z_{\rm eff}^2 g_{\rm eff} {\rm e}^{-E/{\rm k}T}}{E_{\rm keV} T_{\rm keV}^{1/2}}, \tag{31}$$

where in this case $P_{\rm ff}$ is in photons $\times$ m^3 s^{-1} keV^{-1} and $E_{\rm keV}$ is the energy in keV. The total amount of radiation produced by this process is given by

$$W_{\rm tot} = \frac{4.856 \times 10^{-37}\ {\rm W\,m^3}}{\sqrt{T_{\rm keV}}} \int_0^\infty Z_{\rm eff}^2 g_{\rm ff} {\rm e}^{-E/{\rm k}T}\, {\rm d}E_{\rm keV}. \tag{32}$$

From (29) we see immediately that the Bremsstrahlung spectrum (expressed in W m^{-3} keV^{-1}) is flat for $E \ll {\rm k}T$, and for $E > {\rm k}T$ it drops exponentially. In order to measure the temperature of a hot plasma, one needs to measure near $E \simeq {\rm k}T$. The Gaunt factor $g_{\rm ff}$ can be calculated analytically; there are both tables and asymptotic approximations available. In general, $g_{\rm ff}$ depends on both $E/{\rm k}T$ and ${\rm k}T/Z_{\rm eff}$.

For a plasma (29) needs to be summed over all ions that are present in order to calculate the total amount of Bremsstrahlung radiation. For cosmic abundances, hydrogen and helium usually yield the largest contribution. Frequently, one defines an average Gaunt factor $G_{\rm ff}$ by

$$G_{\rm ff} = \sum_i \left(\frac{n_{\rm i}}{n_{\rm e}}\right) Z_{{\rm eff},i}^2 g_{{\rm eff},i}. \tag{33}$$

5.1.2 Free-Bound Emission

Free-bound emission occurs during radiative recombination (Sect. 3.3.1). The energy of the emitted photon is at least the ionisation energy of the recombined ion (for recombination to the ground level) or the ionisation energy that corresponds to the excited state (for recombination to higher levels). From the recombination rate (see Sect. 3.3.1) the free-bound emission is determined immediately:

$$\epsilon_{\rm fb} = \sum_i n_{\rm e} n_{\rm i} R_r. \tag{34}$$

Also here it is possible to define an effective Gaunt factor $G_{\rm fb}$. Free-bound emission is in practise often very important. For example in CIE for ${\rm k}T = 0.1$ keV, free-bound emission is the dominant continuum mechanism for $E > 0.1$ keV; for ${\rm k}T = 1$ keV it dominates above 3 keV. For ${\rm k}T \gg 1$ keV Bremsstrahlung is always the most important mechanism, and for ${\rm k}T \ll 0.1$ keV free-bound emission dominates. See also Fig. 8.

Of course, under conditions of photoionisation equilibrium free-bound emission is even more important, because there are more recombinations than in the CIE case (because T is lower, at comparable ionisation).

5.1.3 Two Photon Emission

This process is in particular important for hydrogen-like or helium-like ions. After a collision with a free electron, an electron from a bound 1s shell is excited to the 2s shell. The quantum-mechanical selection rules do not allow that the 2s electron decays back to the 1s orbit by a

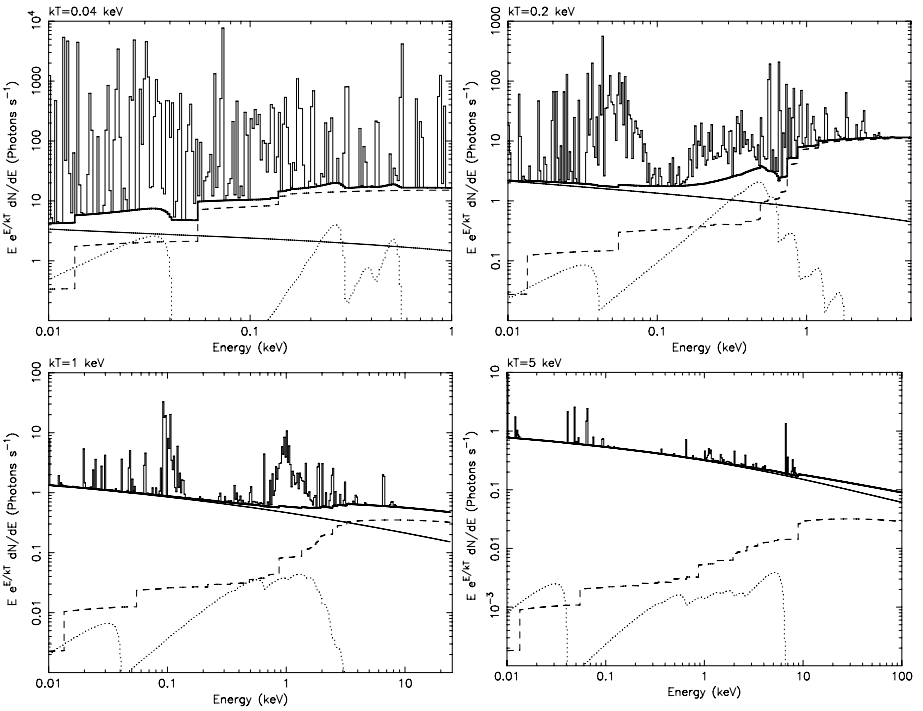


Fig. 8 Emission spectra of plasmas with solar abundances. The histogram indicates the total spectrum, including line radiation. The spectrum has been binned in order to show better the relative importance of line radiation. The *thick solid line* is the total continuum emission, the *thin solid line* the contribution due to Bremsstrahlung, the *dashed line* free-bound emission and the *dotted line* two-photon emission. Note the scaling with $Ee^{E/\mathrm{k}T}$ along the y-axis

radiative transition. Usually the ion will then be excited to a higher level by another collision, for example from 2s to 2p, and then it can decay radiatively back to the ground state (1s). However, if the density is very low ($n_\mathrm{e} \ll n_\mathrm{e,crit}$, (36)–(37)), the probability for a second collision is very small and in that case two-photon emission can occur: the electron decays from the 2s orbit to the 1s orbit while emitting *two* photons. Energy conservation implies that the total energy of both photons should equal the energy difference between the 2s and 1s level ($E_\mathrm{2phot} = E_\mathrm{1s} - E_\mathrm{2s}$). From symmetry considerations it is clear that the spectrum must be symmetrical around $E = 0.5E_\mathrm{2phot}$, and further that it should be zero for $E = 0$ and $E = E_\mathrm{2phot}$. An empirical approximation for the shape of the spectrum is given by:

$$F(E) \sim \sqrt{\sin(\pi E/E_\mathrm{2phot})}. \tag{35}$$

An approximation for the critical density below which two photon emission is important can be obtained from a comparison of the radiative and collisional rates from the upper (2s) level, and is given by (Mewe et al. 1986):

$$\text{H-like}: \quad n_\mathrm{e,crit} = 7 \times 10^9 \ \mathrm{m}^{-3} Z^{9.5} \tag{36}$$

$$\text{He-like}: \quad n_\mathrm{e,crit} = 2 \times 10^{11} \ \mathrm{m}^{-3} (Z-1)^{9.5}. \tag{37}$$

For example for carbon two photon emission is important for densities below 10^{17} m^{-3}, which is the case for many astrophysical applications. Also in this case one can determine an average Gaunt factor $G_{2\mathrm{phot}}$ by averaging over all ions. Two photon emission is important in practise for $0.5 \lesssim kT \lesssim 5$ keV, and then in particular for the contributions of C, N and O between 0.2 and 0.6 keV. See also Fig. 8.

5.2 Line Emission Processes

Apart from continuum radiation, line radiation plays an important role for thermal plasmas. In some cases the flux, integrated over a broad energy band, can be completely dominated by line radiation (see Fig. 8). The production process of line radiation can be broken down in two basic steps: the excitation and the emission process.

5.2.1 Excitation Process

An atom or ion must first be brought into an excited state before it can emit line photons. There are several physical processes that can contribute to this.

The most important process is usually collisional excitation (Sect. 3.1), in particular for plasmas in CIE. The collision of an electron with the ion brings it in an excited state.

A second way to excite the ion is by absorbing a photon with the proper energy. We discuss this process in more detail in Sect. 6.

Alternatively, inner shell ionisation (either by the collision with a free electron, Sect. 3.2.1 or by the photoelectric effect, Sect. 3.2.2) brings the ion in an excited state.

Finally, the ion can also be brought in an excited state by capturing a free electron in one of the free energy levels above the ground state (radiative recombination, Sect. 3.3.1), or through dielectronic recombination (Sect. 3.3.2).

5.2.2 Line Emission

It does not matter by whatever process the ion is brought into an excited state j, whenever it is in such a state it may decay back to the ground state or any other lower energy level i by emitting a photon. The probability per unit time that this occurs is given by the spontaneous transition probability A_{ij} (units: s^{-1}) which is a number that is different for each transition. The total line power P_{ij} (photons per unit time and volume) is then given by

$$P_{ij} = A_{ij} n_j \tag{38}$$

where n_j is the number density of ions in the excited state j. For the most simple case of excitation from the ground state g (rate S_{gj}) followed by spontaneous emission, one can simply approximate $n_g n_e S_{gj} = n_j A_{gj}$. From this equation, the relative population $n_j/n_g \ll 1$ is determined, and then using (38) the line flux is determined. In realistic situations, however, things are more complicated. First, the excited state may also decay to other intermediate states if present, and also excitations or cascades from other levels may play a role. Furthermore, for high densities also collisional excitation or de-excitation to and from other levels becomes important. In general, one has to solve a set of equations for all energy levels of an ion where all relevant population and depopulation processes for that level are taken into account. For the resulting solution vector n_j, the emitted line power is then simply given by (38).

Note that not all possible transitions between the different energy levels are allowed. There are strict quantum mechanical selection rules that govern which lines are allowed; see

for instance Herzberg (1944) or Mewe (1999). Sometimes there are higher order processes that still allow a forbidden transition to occur, albeit with much smaller transition probabilities A_{ij}. But if the excited state j has no other (fast enough) way to decay, these forbidden lines occur and the lines can be quite strong, as their line power is essentially governed by the rate at which the ion is brought into its excited state j.

One of the most well known groups of lines is the He-like 1s–2p triplet. Usually the strongest line is the resonance line, an allowed transition. The forbidden line has a similar strength as the resonance line, for the low density conditions in the ISM and intracluster medium, but it can be relatively enhanced in recombining plasmas, or relatively reduced in high density plasmas like stellar coronal loops. In between both lines is the intercombination line. In fact, this intercombination line is a doublet but for the lighter elements both components cannot be resolved. But see Fig. 6 for the case of iron.

5.2.3 Line Width

For most X-ray spectral lines, the line profile of a line with energy E can be approximated with a Gaussian $\exp(-\Delta E^2/2\sigma^2)$ with σ given by $\sigma/E = \sigma_\mathrm{v}/c$ where the velocity dispersion is

$$\sigma_\mathrm{v}^2 = \sigma_\mathrm{t}^2 + \mathrm{k}T_\mathrm{i}/m_\mathrm{i}. \tag{39}$$

Here T_i is the ion temperature (not necessarily the same as the electron temperature), and σ_t is the root mean squared turbulent velocity of the emitting medium. For large ion temperature, turbulent velocity or high spectral resolution this line width can be measured, but in most cases the lines are not resolved for CCD type spectra.

5.2.4 Resonance Scattering

Resonance scattering is a process where a photon is absorbed by an atom and then re-emitted as a line photon of the same energy into a different direction. As for strong resonance lines (allowed transitions) the transition probabilities A_{ij} are large, the time interval between absorption and emission is extremely short, and that is the reason why the process effectively can be regarded as a scattering process. We discuss the absorption properties in Sect. 6.3, and have already discussed the spontaneous emission in Sect. 5.2.2.

Resonance scattering of X-ray photons is potentially important in the dense cores of some clusters of galaxies for a few transitions of the most abundant elements, as first shown by Gil'fanov et al. (1987). The optical depth for scattering can be written conveniently as (cf. also Sect. 6.3):

$$\tau = \frac{4240 f N_{24} (\frac{n_\mathrm{i}}{n_\mathrm{Z}})(\frac{n_\mathrm{Z}}{n_\mathrm{H}})(\frac{M}{T_\mathrm{keV}})^{1/2}}{E_\mathrm{keV}\{1 + \frac{0.0522 M v_{100}^2}{T_\mathrm{keV}}\}^{1/2}}, \tag{40}$$

where f is the absorption oscillator strength of the line (of order unity for the strongest spectral lines), E_keV the energy in keV, N_{24} the hydrogen column density in units of 10^{24} m^{-2}, n_i the number density of the ion, n_Z the number density of the element, M the atomic weight of the ion, T_keV the ion temperature in keV (assumed to be equal to the electron temperature) and v_{100} the micro-turbulence velocity in units of 100 km/s. Resonance scattering in clusters causes the radial intensity profile on the sky of an emission line to become weaker in the cluster core and stronger in the outskirts, without destroying photons. By comparing the radial line profiles of lines with different optical depth, for instance the 1s−2p and 1s−3p lines

of O VII or Fe XXV, one can determine the optical depth and hence constrain the amount of turbulence in the cluster.

Another important application was proposed by Churazov et al. (2001). They show that for WHIM filaments the resonance line of the O VII triplet can be enhanced significantly compared to the thermal emission of the filament due to resonant scattering of X-ray background photons on the filament. The ratio of this resonant line to the other lines of the triplet therefore can be used to estimate the column density of a filament.

5.3 Some Important Line Transitions

In Tables 4 and 5 we list the 100 strongest emission lines under CIE conditions. Note that each line has its peak emissivity at a different temperature. In particular some of the H-like and He-like transitions are strong, and further the so-called Fe-L complex (lines from $n = 2$ in Li-like to Ne-like iron ions) is prominent. At longer wavelengths, the L-complex of Ne Mg, Si and S gives strong soft X-ray lines. At very short wavelengths, there are not many strong emission lines: between 6–7 keV, the Fe-K emission lines are the strongest spectral features.

6 Absorption Processes

6.1 Continuum Versus Line Absorption

X-rays emitted by cosmic sources do not travel unattenuated to a distant observer. This is because intervening matter in the line of sight absorbs a part of the X-rays. With low-resolution instruments, absorption can be studied only through the measurement of broad-band flux depressions caused by continuum absorption. However, at high spectral resolution also absorption lines can be studied, and in fact absorption lines offer more sensitive tools to detect weak intervening absorption systems. We illustrate this in Fig. 9. At an O VIII column density of 10^{21} m^{-2}, the absorption edge has an optical depth of 1%; for the same column density, the core of the Lyα line is already saturated and even for 100 times lower column density, the line core still has an optical depth of 5%.

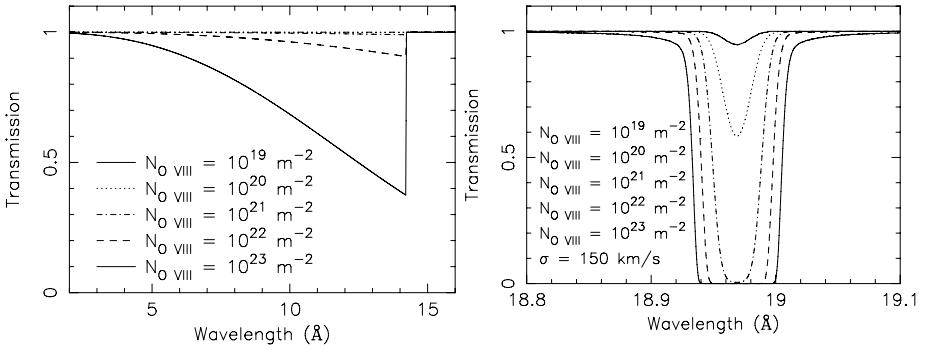


Fig. 9 Continuum (*left*) and Lyα (*right*) absorption spectrum for a layer consisting of O VIII ions with column densities as indicated

Table 4 The strongest emission lines for a plasma with proto-solar abundances (Lodders 2003) in the X-ray band 43 Å $< \lambda <$ 100 Å. At longer wavelengths sometimes a few lines from the same multiplet have been added. All lines include unresolved dielectronic satellites. $T_{\max}$ (K) is the temperature at which the emissivity peaks, $Q_{\max} = P/(n_e n_H)$, with P the power per unit volume at $T_{\max}$, and $Q_{\max}$ is in units of 10^{-36} W m^3

E (eV)	λ (Å)	$-\log Q_{\max}$	$\log T_{\max}$	Ion	Iso-el. seq.	Lower level	Upper level
126.18	98.260	1.35	5.82	Ne VIII	Li	2p $^2P_{3/2}$	3d $^2D_{5/2}$
126.37	98.115	1.65	5.82	Ne VIII	Li	2p $^2P_{1/2}$	3d $^2D_{3/2}$
127.16	97.502	1.29	5.75	Ne VII	Be	2s $^1S_0$	3p $^1P_1$
128.57	96.437	1.61	5.47	Si V	Ne	2p $^1S_0$	3d $^1P_1$
129.85	95.483	1.14	5.73	Mg VI	N	2p $^4S_{3/2}$	3d $^4P_{5/2,3/2,1/2}$
132.01	93.923	1.46	6.80	Fe XVIII	F	2s $^2P_{3/2}$	2p $^2S_{1/2}$
140.68	88.130	1.21	5.75	Ne VII	Be	2p $^3P_1$	4d $^3D_{2,3}$
140.74	88.092	1.40	5.82	Ne VIII	Li	2s $^2S_{1/2}$	3p $^2P_{3/2}$
147.67	83.959	0.98	5.86	Mg VII	C	2p ^3P	3d ^3D,^1D,^3F
148.01	83.766	1.77	5.86	Mg VII	C	2p $^3P_2$	3d $^3P_2$
148.56	83.457	1.41	5.90	Fe IX	Ar	3p $^1S_0$	4d $^3P_1$
149.15	83.128	1.69	5.69	Si VI	F	2p $^2P_{3/2}$	(^{3}P)3d $^2D_{5/2}$
149.38	83.000	1.48	5.74	Mg VI	N	2p $^4S_{3/2}$	4d $^4P_{5/2,3/2,1/2}$
150.41	82.430	1.49	5.91	Fe IX	Ar	3p $^1S_0$	4d $^1P_1$
154.02	80.501	1.75	5.70	Si VI	F	2p $^2P_{3/2}$	(^{1}D) 3d $^2D_{5/2}$
159.23	77.865	1.71	6.02	Fe X	Cl	3p $^2P_{1/2}$	4d $^2D_{5/2}$
165.24	75.034	1.29	5.94	Mg VIII	B	2p $^2P_{3/2}$	3d $^2D_{5/2}$
165.63	74.854	1.29	5.94	Mg VIII	B	2p $^2P_{1/2}$	3d $^2D_{3/2}$
170.63	72.663	1.07	5.76	S VII	Ne	2p $^1S_0$	3s $^3P_{1,2}$
170.69	72.635	1.61	6.09	Fe XI	S	3p $^3P_2$	4d $^3D_3$
171.46	72.311	1.56	6.00	Mg IX	Be	2p $^1P_1$	3d $^1D_2$
171.80	72.166	1.68	6.08	Fe XI	S	3p $^1D_2$	4d $^1F_3$
172.13	72.030	1.44	6.00	Mg IX	Be	2p $^3P_{2,1}$	3s $^3S_1$
172.14	72.027	1.40	5.76	S VII	Ne	2p $^1S_0$	3s $^1P_1$
177.07	70.020	1.18	5.84	Si VII	O	2p ^3P	3d ^3D,^3P
177.98	69.660	1.36	6.34	Fe XV	Mg	3p $^1P_1$	4s $^1S_0$
177.99	69.658	1.57	5.96	Si VIII	N	2p $^4S_{3/2}$	3s $^4P_{5/2,3/2,1/2}$
179.17	69.200	1.61	5.71	Si VI	F	2p ^2P	4d ^2P, ^2D
186.93	66.326	1.72	6.46	Fe XVI	Na	3d ^2D	4f ^2F
194.58	63.719	1.70	6.45	Fe XVI	Na	3p $^2P_{3/2}$	4s $^2S_{1/2}$
195.89	63.294	1.64	6.08	Mg X	Li	2p $^2P_{3/2}$	3d $^2D_{5/2}$
197.57	62.755	1.46	6.00	Mg IX	Be	2s $^1S_0$	3p $^1P_1$
197.75	62.699	1.21	6.22	Fe XIII	Si	3p $^3P_1$	4d $^3D_2$
198.84	62.354	1.17	6.22	Fe XIII	Si	3p $^3P_0$	4d $^3D_1$
199.65	62.100	1.46	6.22	Fe XIII	Si	3p $^3P_1$	4d $^3P_0$
200.49	61.841	1.29	6.07	Si IX	C	2p $^3P_2$	3s $^3P_1$
203.09	61.050	1.06	5.96	Si VIII	N	2p $^4S_{3/2}$	3d $^4P_{5/2,3/2,1/2}$
203.90	60.807	1.69	5.79	S VII	Ne	2p $^1S_0$	3d $^3D_1$
204.56	60.610	1.30	5.79	S VII	Ne	2p $^1S_0$	3d $^1P_1$
223.98	55.356	1.00	6.08	Si IX	C	2p ^3P	3d ^3D,^1D,^3F

Table 4 (*Continued*)

E (eV)	λ (Å)	−log Q_{max}	log T_{max}	Ion	Iso-el. seq.	Lower level	Upper level
234.33	52.911	1.34	6.34	Fe XV	Mg	$3s\ ^1S_0$	$4p\ ^1P_1$
237.06	52.300	1.61	6.22	Si XI	Be	$2p\ ^1P_1$	$3s\ ^1S_0$
238.43	52.000	1.44	5.97	Si VIII	N	$2p\ ^4S_{3/2}$	$4d\ ^4P_{5/2,3/2,1/2}$
244.59	50.690	1.30	6.16	Si X	B	$2p\ ^2P_{3/2}$	$3d\ ^2D_{5/2}$
245.37	50.530	1.30	6.16	Si X	B	$2p\ ^2P_{1/2}$	$3d\ ^2D_{3/2}$
251.90	49.220	1.45	6.22	Si XI	Be	$2p\ ^1P_1$	$3d\ ^1D_2$
252.10	49.180	1.64	5.97	Ar IX	Ne	$2p\ ^1S_0$	$3s\ ^3P_{1,2}$
261.02	47.500	1.47	6.06	S IX	O	$2p\ ^3P$	$3d\ ^3D,^3P$
280.73	44.165	1.60	6.30	Si XII	Li	$2p\ ^2P_{3/2}$	$3d\ ^2D_{5/2}$
283.46	43.740	1.46	6.22	Si XI	Be	$2s\ ^1S_0$	$3p\ ^1P_1$

6.2 Continuum Absorption

Continuum absorption can be calculated simply from the photoionisation cross sections, that we discussed already in Sect. 3.2.2. The total continuum opacity $\tau_{\rm cont}$ can be written as

$$\tau_{\rm cont}(E) \equiv N_{\rm H}\sigma_{\rm cont}(E) = \sum_i N_i \sigma_i(E) \tag{41}$$

i.e. by averaging over the various ions i with column density N_i. Accordingly, the continuum transmission $T(E)$ of such a clump of matter can be written as $T(E) = \exp(-\tau_{\rm cont}(E))$. For a worked out example see also Sect. 6.6.

6.3 Line Absorption

When light from a background source shines through a clump of matter, a part of the radiation can be absorbed. We discussed already the continuum absorption. The transmission in a spectral line at wavelength λ is given by

$$T(\lambda) = \mathrm{e}^{-\tau(\lambda)} \tag{42}$$

with

$$\tau(\lambda) = \tau_0 \varphi(\lambda) \tag{43}$$

where $\varphi(\lambda)$ is the line profile and τ_0 is the opacity at the line centre λ_0, given by:

$$\tau_0 = \frac{\alpha \mathrm{h} \lambda f N_i}{2\sqrt{2\pi} m_{\rm e} \sigma_{\rm v}}. \tag{44}$$

Apart from the fine structure constant α and Planck's constant h, the optical depth also depends on the properties of the absorber, namely the ionic column density N_i and the velocity dispersion $\sigma_{\rm v}$. Furthermore, it depends on the oscillator strength f which is a dimensionless quantity that is different for each transition and is of order unity for the strongest transitions.

In the simplest approximation, a Gaussian profile $\varphi(\lambda) = \exp -(\lambda - \lambda_0)^2/b^2)$ can be adopted, corresponding to pure Doppler broadening for a thermal plasma. Here $b = \sqrt{2}\sigma$

Table 5 As Table 4, but for $\lambda < 43$ Å

E (eV)	λ (Å)	$-\log Q_{\max}$	$\log T_{\max}$	Ion	Iso-el. seq.	Lower level	Upper level
291.52	42.530	1.32	6.18	S X	N	2p $^4S_{3/2}$	3d $^4P_{5/2,3/2,1/2}$
298.97	41.470	1.31	5.97	C V	He	1s $^1S_0$	2s $^3S_1$ (f)
303.07	40.910	1.75	6.29	Si XII	Li	2s $^2S_{1/2}$	3p $^2P_{3/2}$
307.88	40.270	1.27	5.98	C V	He	1s $^1S_0$	2p $^1P_1$ (r)
315.48	39.300	1.37	6.28	S XI	C	2p ^3P	3d ^3D,^1D,^3F
336.00	36.900	1.58	6.19	S X	N	2p $^4S_{3/2}$	4d $^4P_{5/2,3/2,1/2}$
339.10	36.563	1.56	6.34	S XII	B	2p $^2P_{3/2}$	3d $^2D_{5/2}$
340.63	36.398	1.56	6.34	S XII	B	2p $^2P_{1/2}$	3d $^2D_{3/2}$
367.47	33.740	1.47	6.13	C VI	H	1s $^2S_{1/2}$	2p $^2P_{1/2}$ (Lyα)
367.53	33.734	1.18	6.13	C VI	H	1s $^2S_{1/2}$	2p $^2P_{3/2}$ (Lyα)
430.65	28.790	1.69	6.17	N VI	He	1s $^1S_0$	2p $^1P_1$ (r)
500.36	24.779	1.68	6.32	N VII	H	1s $^2S_{1/2}$	2p $^2P_{3/2}$ (Lyα)
560.98	22.101	0.86	6.32	O VII	He	1s $^1S_0$	2s $^3S_1$ (f)
568.55	21.807	1.45	6.32	O VII	He	1s $^1S_0$	2p $^3P_{2,1}$ (i)
573.95	21.602	0.71	6.33	O VII	He	1s $^1S_0$	2p $^1P_1$ (r)
653.49	18.973	1.05	6.49	O VIII	H	1s $^2S_{1/2}$	2p $^2P_{1/2}$ (Lyα)
653.68	18.967	0.77	6.48	O VIII	H	1s $^2S_{1/2}$	2p $^2P_{3/2}$ (Lyα)
665.62	18.627	1.58	6.34	O VII	He	1s $^1S_0$	3p $^1P_1$
725.05	17.100	0.87	6.73	Fe XVII	Ne	2p $^1S_0$	3s $^3P_2$
726.97	17.055	0.79	6.73	Fe XVII	Ne	2p $^1S_0$	3s $^3P_1$
738.88	16.780	0.87	6.73	Fe XVII	Ne	2p $^1S_0$	3s $^1P_1$
771.14	16.078	1.37	6.84	Fe XVIII	F	2p $^2P_{3/2}$	3s $^4P_{5/2}$
774.61	16.006	1.55	6.50	O VIII	H	1s $^2S_{1/2}$	3p $^2P_{1/2,3/2}$ (Lyβ)
812.21	15.265	1.12	6.74	Fe XVII	Ne	2p $^1S_0$	3d $^3D_1$
825.79	15.014	0.58	6.74	Fe XVII	Ne	2p $^1S_0$	3d $^1P_1$
862.32	14.378	1.69	6.84	Fe XVIII	F	2p $^2P_{3/2}$	3d $^2D_{5/2}$
872.39	14.212	1.54	6.84	Fe XVIII	F	2p $^2P_{3/2}$	3d $^2S_{1/2}$
872.88	14.204	1.26	6.84	Fe XVIII	F	2p $^2P_{3/2}$	3d $^2D_{5/2}$
896.75	13.826	1.66	6.76	Fe XVII	Ne	2s $^1S_0$	3p $^1P_1$
904.99	13.700	1.61	6.59	Ne IX	He	1s $^1S_0$	2s $^3S_1$ (f)
905.08	13.699	1.61	6.59	Ne IX	He	1s $^1S_0$	2s $^3S_1$ (f)
916.98	13.521	1.35	6.91	Fe XIX	O	2p $^3P_2$	3d $^3D_3$
917.93	13.507	1.68	6.91	Fe XIX	O	2p $^3P_2$	3d $^3P_2$
922.02	13.447	1.44	6.59	Ne IX	He	1s $^1S_0$	2p $^1P_1$ (r)
965.08	12.847	1.51	6.98	Fe XX	N	2p $^4S_{3/2}$	3d $^4P_{5/2}$
966.59	12.827	1.44	6.98	Fe XX	N	2p $^4S_{3/2}$	3d $^4P_{3/2}$
1009.2	12.286	1.12	7.04	Fe XXI	C	2p $^3P_0$	3d $^3D_1$
1011.0	12.264	1.46	6.73	Fe XVII	Ne	2p $^1S_0$	4d $^3D_1$
1021.5	12.137	1.77	6.77	Ne X	H	1s $^2S_{1/2}$	2p $^2P_{1/2}$ (Lyα)
1022.0	12.132	1.49	6.76	Ne X	H	1s $^2S_{1/2}$	2p $^2P_{3/2}$ (Lyα)
1022.6	12.124	1.39	6.73	Fe XVII	Ne	2p $^1S_0$	4d $^1P_1$
1053.4	11.770	1.38	7.10	Fe XXII	B	2p $^2P_{1/2}$	3d $^2D_{3/2}$

Table 5 (*Continued*)

E (eV)	λ (Å)	−log Q_{max}	log T_{max}	Ion	Iso-el. seq.	Lower level	Upper level
1056.0	11.741	1.51	7.18	Fe XXIII	Be	$2p\ ^1P_1$	$3d\ ^1D_2$
1102.0	11.251	1.73	6.74	Fe XVII	Ne	$2p\ ^1S_0$	$5d\ ^3D_1$
1352.1	9.170	1.66	6.81	Mg XI	He	$1s\ ^1S_0$	$2p\ ^1P_1$ (r)
1472.7	8.419	1.76	7.00	Mg XII	H	$1s\ ^2S_{1/2}$	$2p\ ^2P_{3/2}$ (Lyα)
1864.9	6.648	1.59	7.01	Si XIII	He	$1s\ ^1S_0$	$2p\ ^1P_1$ (r)
2005.9	6.181	1.72	7.21	Si XIV	H	$1s\ ^2S_{1/2}$	$2p\ ^2P_{3/2}$ (Lyα)
6698.6	1.851	1.43	7.84	Fe XXV	He	$1s\ ^1S_0$	$2p\ ^1P_1$ (r)
6973.1	1.778	1.66	8.17	Fe XXVI	H	$1s\ ^2S_{1/2}$	$2p\ ^2P_{3/2}$ (Lyα)

with σ the normal Gaussian root-mean-square width. The full width at half maximum of this profile is given by $\sqrt{\ln 256}\sigma$ or approximately 2.35σ. One may even include a turbulent velocity σ_t into the velocity width σ_v, such that

$$\sigma_v^2 = \sigma_t^2 + kT/m_i \tag{45}$$

with m_i the mass of the ion (we have tacitly changed here from wavelength to velocity units through the scaling $\Delta\lambda/\lambda_0 = \Delta v/c$).

The equivalent width W of the line is calculated from

$$W = \frac{\lambda\sigma}{c} \int_{-\infty}^{\infty} \left[1 - \exp\left(-\tau_0 e^{-y^2/2}\right)\right] dy. \tag{46}$$

For the simple case that $\tau_0 \ll 1$, the integral can be evaluated as

$$\tau_0 \ll 1: \quad W = \frac{\alpha h \lambda^2 f N_i}{2 m_e c} = \frac{1}{2}\alpha \frac{h\nu}{m_e c^2} f \lambda^2 N_i. \tag{47}$$

A Gaussian line profile is only a good approximation when the Doppler width of the line is larger than the natural width of the line. The natural line profile for an absorption line is a Lorentz profile $\varphi(\lambda) = 1/(1 + x^2)$ with $x = 4\pi\Delta\nu/A$. Here $\Delta\nu$ is the frequency difference $\nu - \nu_0$ and A is the total transition probability from the upper level downwards to any level, including all radiative and Auger decay channels.

Convolving the intrinsic Lorentz profile with the Gaussian Doppler profile due to the thermal or turbulent motion of the plasma, gives the well-known Voigt line profile

$$\varphi = H(a, y) \tag{48}$$

where

$$a = A\lambda/4\pi b \tag{49}$$

and $y = c\Delta\lambda/b\lambda$. The dimensionless parameter a (not to be confused with the a in (4)) represents the relative importance of the Lorentzian term ($\sim A$) to the Gaussian term ($\sim b$). It should be noted here that formally for $a > 0$ the parameter τ_0 defined by (44) is not exactly the optical depth at line centre, but as long as $a \ll 1$ it is a fair approximation. For $a \gg 1$ it can be shown that $H(a, 0) \to 1/a\sqrt{\pi}$.

6.4 Some Important X-ray Absorption Lines

There are several types of absorption lines. The most well-known are the normal strong resonance lines, which involve electrons from the outermost atomic shells. Examples are the 1s–2p line of O VII at 21.60 Å, the O VIII Lyα doublet at 18.97 Å and the well known 2s–2p doublet of O VI at 1032 and 1038 Å. See Richter et al. 2008—Chap. 3, this issue, for an extensive discussion on these absorption lines in the WHIM.

The other class are the inner-shell absorption lines. In this case the excited atom is often in an unstable state, with a large probability for the Auger process (Sect. 3.2.4). As a result, the parameter a entering the Voigt profile (49) is large and therefore the lines can have strong damping wings.

In some cases the lines are isolated and well resolved, like the O VII and O VIII 1s–2p lines mentioned above. However, in other cases the lines may be strongly blended, like for the higher principal quantum number transitions of any ion in case of high column densities. Another important group of strongly blended lines are the inner-shell transitions of heavier elements like iron. They form so-called unresolved transition arrays (UTAs); the individual lines are no longer recognisable. The first detection of these UTAs in astrophysical sources was in a quasar (Sako et al. 2001).

In Table 6 we list the 70 most important X-ray absorption lines for $\lambda < 100$ Å. The importance was determined by calculating the maximum ratio W/λ that these lines reach for any given temperature.

Note the dominance of lines from all ionisation stages of oxygen in the 17–24 Å band. Furthermore, the Fe IX and Fe XVII lines are the strongest iron features. These lines are weaker than the oxygen lines because of the lower abundance of iron; they are stronger than their neighbouring iron ions because of somewhat higher oscillator strengths and a relatively higher ion fraction (cf. Fig. 7).

Note that the strength of these lines can depend on several other parameters, like turbulent broadening, deviations from CIE, saturation for higher column densities, etc., so some care should be taken when the strongest line for other physical conditions are sought.

6.5 Curve of Growth

In most practical situations, X-ray absorption lines are unresolved or poorly resolved. As a consequence, the physical information about the source can only be retrieved from basic parameters such as the line centroid (for Doppler shifts) and equivalent width W (as a measure of the line strength).

As outlined in Sect. 6.3, the equivalent width W of a spectral line from an ion i depends, apart from atomic parameters, only on the column density N_i and the velocity broadening σ_v. A useful diagram is the so-called curve of growth. Here we present it in the form of a curve giving W versus N_i for a fixed value of σ_v.

The curve of growth has three regimes. For low column densities, the optical depth at line centre τ_0 is small, and W increases linearly with N_i, independent of σ_v (47). For increasing column density, τ_0 starts to exceed unity and then W increases only logarithmically with N_i. Where this happens exactly, depends on σ_v. In this intermediate logarithmic branch W depends both on σ_v and N_i, and hence by comparing the measured W with other lines from the same ion, both parameters can be estimated. At even higher column densities, the effective opacity increases faster (proportional to $\sqrt{N_i}$), because of the Lorentzian line wings. Where this happens depends on the Voigt parameter a (49). In this range of densities, W does not depend any more on σ_v.

Table 6 The most important X-ray absorption lines with $\lambda < 100$ Å. Calculations are done for a plasma in CIE with proto-solar abundances (Lodders 2003) and only thermal broadening. The calculations are done for a hydrogen column density of 10^{24} m^{-2}, and we list equivalent widths for the temperature $T_{\max}$ (in K) where it reaches a maximum. Equivalent widths are calculated for the listed line only, not taking into account blending by other lines. For saturated, overlapping lines, like the O VIII Lyα doublet near 18.97 Å, the combined equivalent width can be smaller than the sum of the individual equivalent widths

λ (Å)	Ion	$W_{\max}$ (mÅ)	log $T_{\max}$	λ (Å)	Ion	$W_{\max}$ (mÅ)	log $T_{\max}$
9.169	Mg XI	1.7	6.52	31.287	N I	5.5	4.04
13.447	Ne IX	4.6	6.38	33.426	C V	7.0	5.76
13.814	Ne VII	2.2	5.72	33.734	C VI	12.9	6.05
15.014	Fe XVII	5.4	6.67	33.740	C VI	10.3	6.03
15.265	Fe XVII	2.2	6.63	34.973	C V	10.3	5.82
15.316	Fe XV	2.3	6.32	39.240	S XI	6.5	6.25
16.006	O VIII	2.6	6.37	40.268	C V	17.0	5.89
16.510	Fe IX	5.6	5.81	40.940	C IV	7.2	5.02
16.773	Fe IX	3.0	5.81	41.420	C IV	14.9	5.03
17.396	O VII	2.8	6.06	42.543	S X	6.3	6.14
17.768	O VII	4.4	6.11	50.524	SI X	11.0	6.14
18.629	O VII	6.9	6.19	51.807	S VII	7.6	5.68
18.967	O VIII	8.7	6.41	55.305	SI IX	14.0	6.06
18.973	O VIII	6.5	6.39	60.161	S VII	12.2	5.71
19.924	O V	3.2	5.41	61.019	SI VIII	11.9	5.92
21.602	O VII	12.1	6.27	61.070	SI VIII	13.2	5.93
22.006	O VI	5.1	5.48	62.751	Mg IX	11.1	5.99
22.008	O VI	4.1	5.48	68.148	SI VII	10.1	5.78
22.370	O V	5.4	5.41	69.664	SI VII	10.9	5.78
22.571	O IV	3.6	5.22	70.027	SI VII	11.3	5.78
22.739	O IV	5.1	5.24	73.123	SI VII	10.7	5.78
22.741	O IV	7.2	5.23	74.858	Mg VIII	17.1	5.93
22.777	O IV	13.8	5.22	80.449	SI VI	13.0	5.62
22.978	O III	9.4	4.95	80.577	SI VI	12.7	5.61
23.049	O III	7.1	4.96	82.430	Fe IX	14.0	5.91
23.109	O III	15.1	4.95	83.128	SI VI	14.2	5.63
23.350	O II	7.5	4.48	83.511	Mg VII	14.2	5.82
23.351	O II	12.6	4.48	83.910	Mg VII	19.8	5.85
23.352	O II	16.5	4.48	88.079	Ne VIII	15.2	5.80
23.510	O I	7.2	4.00	95.385	Mg VI	15.5	5.67
23.511	O I	15.0	4.00	95.421	Mg VI	18.7	5.69
24.779	N VII	4.6	6.20	95.483	Mg VI	20.5	5.70
24.900	N VI	4.0	5.90	96.440	SI V	16.8	5.33
28.465	C VI	4.6	6.01	97.495	Ne VII	22.8	5.74
28.787	N VI	9.8	6.01	98.131	Ne VI	16.7	5.65

In Fig. 10 we show a few characteristic examples. The examples of O I and O VI illustrate the higher W for UV lines as compared to X-ray lines (cf. (47)) as well as the fact that inner shell transitions (in this case the X-ray lines) have larger a-values and hence reach sooner the

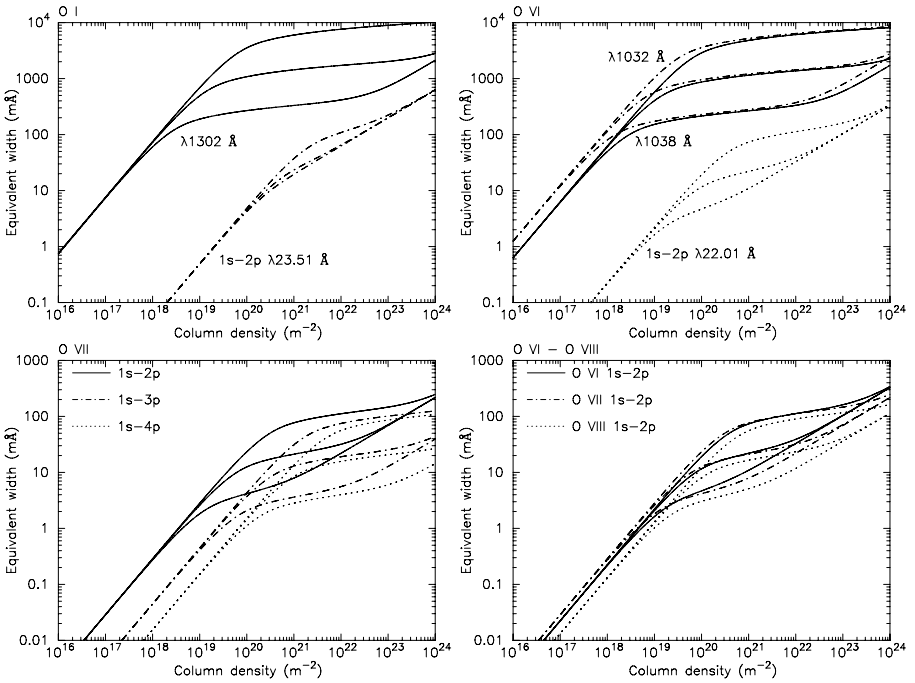


Fig. 10 Equivalent width versus column density for a few selected oxygen absorption lines. The curves have been calculated for Gaussian velocity dispersions $\sigma = b/\sqrt{2}$ of 10, 50 and 250 km s^{-1}, from *bottom* to *top* for each line. Different spectral lines are indicated with different line styles

square root branch. The example of O VII shows how different lines from the same ion have different equivalent width ratios depending on σ_v, hence offer an opportunity to determine that parameter. The last frame shows the (non)similarities for similar transitions in different ions. The innershell nature of the transition in O VI yields a higher a value and hence an earlier onset of the square root branch.

6.6 Galactic Foreground Absorption

All radiation from X-ray sources external to our own Galaxy has to pass through the interstellar medium of our Galaxy, and the intensity is reduced by a factor of $e^{-\tau(E)}$ with the optical depth $\tau(E) = \sum_i \sigma_i(E) \int n_i(\ell) d\ell$, with the summation over all relevant ions i and the integration over the line of sight $d\ell$. The absorption cross section $\sigma_i(E)$ is often taken to be simply the (continuum) photoionisation cross section, but for high-resolution spectra it is necessary to include also the line opacity, and for very large column densities also other processes such as Compton ionisation or Thomson scattering.

For a cool, neutral plasma with cosmic abundances one often writes

$$\tau = \sigma_{\rm eff}(E) N_{\rm H} \tag{50}$$

where the hydrogen column density $N_{\rm H} \equiv \int n_{\rm H} dx$. In $\sigma_{\rm eff}(E)$ all contributions to the absorption from all elements as well as their abundances are taken into account.

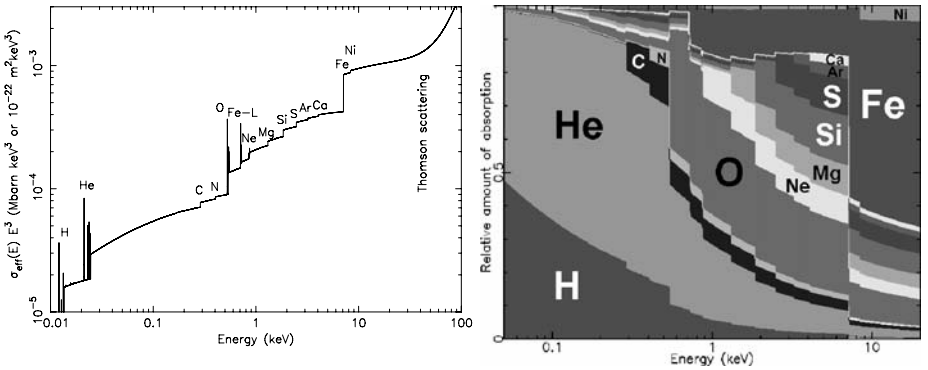


Fig. 11 *Left panel*: Neutral interstellar absorption cross section per hydrogen atom, scaled with E^3. The most important edges with associated absorption lines are indicated, as well as the onset of Thompson scattering around 30 keV. *Right panel*: Contribution of the various elements to the total absorption cross section as a function of energy, for Solar abundances

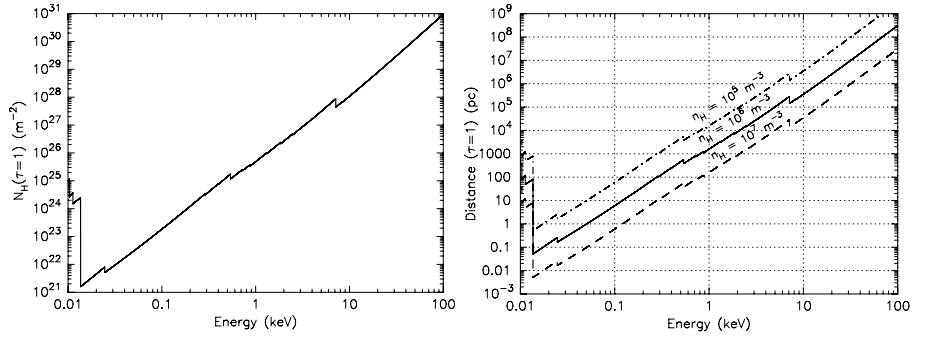


Fig. 12 *Left panel*: Column density for which the optical depth becomes unity. *Right panel*: Distance for which $\tau = 1$ for three characteristic densities

The relative contribution of the elements is also made clear in Fig. 11. Below 0.28 keV (the carbon edge) hydrogen and helium dominate, while above 0.5 keV in particular oxygen is important. At the highest energies, above 7.1 keV, iron is the main opacity source.

Yet another way to represent these data is to plot the column density or distance for which the optical depth becomes unity (Fig. 12). This figure shows that in particular at the lowest energies X-rays are most strongly absorbed. The visibility in that region is thus very limited. Below 0.2 keV it is extremely hard to look outside our Milky Way.

The interstellar medium is by no means a homogeneous, neutral, atomic gas. In fact, it is a collection of regions all with different physical state and composition. This affects its X-ray opacity. We briefly discuss here some of the most important features.

The ISM contains cold gas (<50 K), warm, neutral or lowly ionised gas (6 000–10 000 K) as well as hotter gas (a few million K). We have seen before (see Sect. 3.2.2) that for ions the absorption edges shift to higher energies for higher ionisation. Thus, with instruments of sufficient spectral resolution, the degree of ionisation of the ISM can be deduced from the relative intensities of the absorption edges. Cases with such high column densities are rare, however (and only occur in some AGN outflows), and for the bulk of the ISM the column density of the ionised ISM is low enough that only the narrow absorption lines are visible

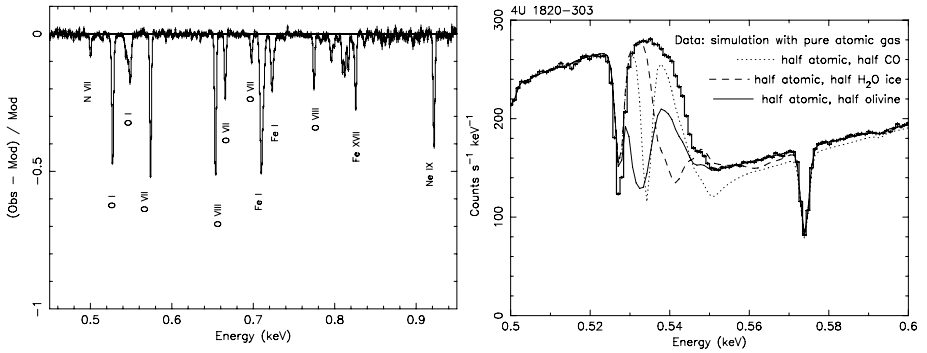


Fig. 13 *Left panel*: Simulated 100 ks absorption spectrum as observed with the Explorer of Diffuse Emission and Gamma-ray Burst Explosions (EDGE), a mission proposed for ESA's Cosmic Vision program. The parameters of the simulated source are similar to those of 4U 1820-303 (Yao and Wang 2006). The plot shows the residuals of the simulated spectrum if the absorption lines in the model are ignored. Several characteristic absorption features of both neutral and ionised gas are indicated.
Right panel: Simulated spectrum for the X-ray binary 4U 1820-303 for 100 ks with the WFS instrument of EDGE. The simulation was done for all absorbing oxygen in its pure atomic state, the models plotted with different line styles show cases where half of the oxygen is bound in CO, water ice or olivine. Note the effective shift of the absorption edge and the different fine structure near the edge. All of this is well resolved by EDGE, allowing a determination of the molecular composition of dust in the line of sight towards this source. The absorption line at 0.574 is due to highly ionised O VII

(see Fig. 9). These lines are only visible when high spectral resolution is used (see Fig. 13). It is important to recognise these lines, as they should not be confused with absorption line from within the source itself. Fortunately, with high spectral resolution the cosmologically redshifted absorption lines from the X-ray source are separated well from the foreground hot ISM absorption lines. Only for lines from the Local Group this is not possible, and the situation is here complicated as the expected temperature range for the diffuse gas within the Local Group is similar to the temperature of the hot ISM.

Another ISM component is dust. A significant fraction of some atoms can be bound in dust grains with varying sizes, as shown below (from Wilms et al. 2000):

H	He	C	N	O	Ne	Mg	Si	S	Ar	Ca	Fe	Ni
0	0	0.5	0	0.4	0	0.8	0.9	0.4	0	0.997	0.7	0.96

The numbers represent the fraction of the atoms that are bound in dust grains. Noble gases like Ne and Ar are chemically inert hence are generally not bound in dust grains, but other elements like Ca exist predominantly in dust grains. Dust has significantly different X-ray properties compared to gas or hot plasma. First, due to the chemical binding, energy levels are broadened significantly or even absent. For example, for oxygen in most bound forms (like H_2O) the remaining two vacancies in the 2p shell are effectively filled by the two electrons from the other bound atom(s) in the molecule. Therefore, the strong 1s–2p absorption line at 23.51 Å (527 eV) is not allowed when the oxygen is bound in dust or molecules, because there is no vacancy in the 2p shell. Transitions to higher shells such as the 3p shell are possible, however, but these are blurred significantly and often shifted due to the interactions in the molecule. Each constituent has its own fine structure near the K-edge (Fig. 13, right panel). This fine structure offers therefore the opportunity to study the (true) chemical composition of the dust, but it should be said that the details of the edges in different important compounds are not always (accurately) known, and can differ depending on the state: for

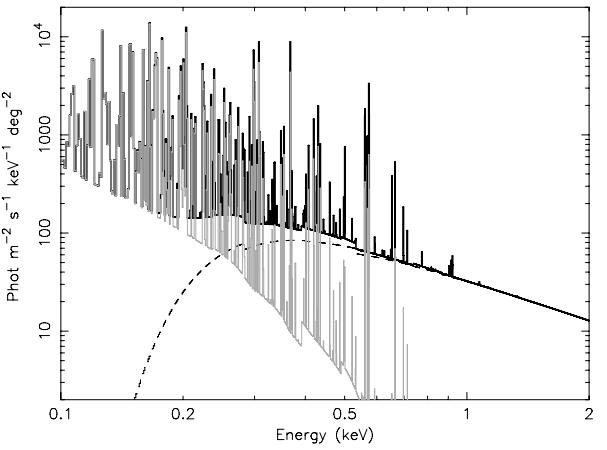


Fig. 14 Typical Cosmic X-ray background spectrum. The *thick solid line* indicates the total spectrum, the *dashed line* the absorbed power law contribution due to unresolved point sources. The contribution of the local hot bubble is indicated by the *thin, grey line*. It dominates the line emission below ∼0.2 keV. The other thermal components are not indicated separately, but they are included in the total; they contribute most of the line flux above 0.2 keV. The calculation is done with a spectral resolution of 1 eV

example water, crystalline and amorphous ice all have different characteristics. On the other hand, the large scale edge structure, in particular when observed at low spectral resolution, is not so much affected. For sufficiently high column densities of dust, self-shielding within the grains should be taken into account, and this reduces the average opacity per atom.

Finally, we mention here that dust also causes scattering of X-rays. This is in particular important for higher column densities. For example, for the Crab nebula ($N_{\rm H} = 3.2 \times 10^{25}$ m^{-2}), at an energy of 1 keV about 10% of all photons are scattered in a halo, of which the widest tails have been observed out to a radius of at least half a degree; the scattered fraction increases with increasing wavelength.

7 Galactic Foreground Emission

The interstellar medium of our Galaxy is a complex medium. While all phases of the ISM can be seen in X-ray absorption towards bright X-ray sources (see previous section), only the hot phase is seen in emission and contributes to the cosmic X-ray background.

This cosmic X-ray background has different components. Kuntz and Snowden (2000) distinguish four different components as outlined below. First, there is an absorbed power-law like component, consisting mostly of unresolved extragalactic point sources. With high spatial resolution like for example available on Chandra, a large part of this component can be resolved into the individual point sources. The second component is the so-called Local Hot Bubble, most likely a large, old supernova remnant embedding our Solar system. The million degree gas of this diffuse medium is almost unabsorbed, as it is nearby and there is little intervening neutral gas. Finally, there are the soft and hard diffuse components, consisting of absorbed, diffuse thermal emission which arises from the disk of our Galaxy but may also contain contributions from a Galactic halo and distant WHIM emission.

Figure 14 shows a simulated spectrum of the total X-ray background. Note however that the details of this spectrum may differ significantly for different parts of the sky. There are variations on large scales, from the Galactic plane to the poles. There are also smaller scale variations. For instance, there are large loops corresponding to old supernova remnants or (super)bubbles, and large molecular cloud complexes obscuring the more distant background components.

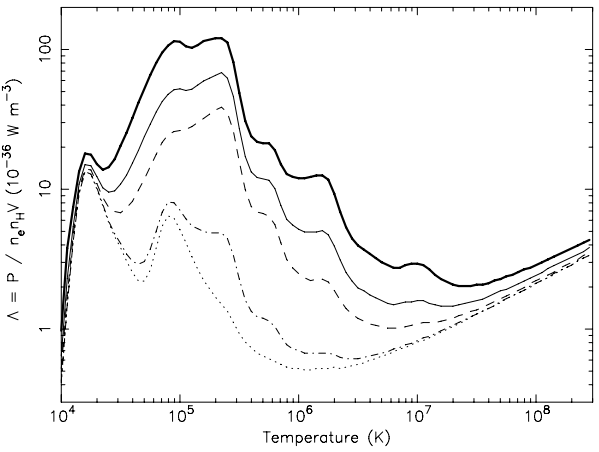


Fig. 15 Cooling curves $\Lambda(T)$ for plasmas of different composition, under CIE conditions. From *top* to *bottom*: 1, 0.3, 0.1, 0.01 and 0.001 times Solar abundances. After Sutherland and Dopita (1993)

8 Cooling Function

The cooling of a plasma by thermal radiation is usually expressed by the cooling function $\Lambda(T)$. This is essentially the bolometric power P emitted as thermal radiation by the plasma, normalised by the emission measure $Y = n_e n_H V$. Accordingly, the characteristic cooling time, defined as the thermal energy of the plasma divided by the emitted power, $t_{cool} = \frac{3}{2}(1 + n_i/n_e)kT/\Lambda n_H$.

Cooling is important for the evolution of several plasmas such as the cool cores of clusters but also for the global evolution of the WHIM. A set of useful cooling curves are given by Sutherland and Dopita (1993). An example is shown in Fig. 15. These are valid under CIE conditions. When photoionisation is important, modifications are needed. Here we have normalised Λ to the emission measure of the plasma, but note that sometimes other scalings are being used (for example expressions containing the ion density n_i or n_e^2). Note that the Solar abundances used in Fig. 15 are the older values of Anders and Grevesse (1989) which differ significantly from more recent estimates, for example Lodders (2003), in particular for important coolants such as iron (new values 26% lower) and oxygen (new values 32% lower). The big hump at low temperatures is produced mainly by line radiation, while the high temperature $\sim T^{0.5}$ tail above 10^7 K is mainly produced by Bremsstrahlung.

Finally, a more recent set of cooling curves, suited in particular for lower temperatures and primordial gas, was published by Maio et al. (2007).

9 Ionisation and Excitation by Non-thermal Electrons

In most cases, astrophysical plasmas are modelled using Maxwellian velocity distributions, even under NEI or PIE conditions. Often this is justified, but there are situations where deviations from a Maxwellian distribution occur. Examples are accelerated electrons penetrating cooler regions, as may occur in or near shocks or flares. We refrain here from a full discussion, but only mention here that spectral codes such as SPEX (Sect. 10.1) allow the calculation of the ionisation balance and the emanating spectrum under such circumstances, provided the true electron distribution can be approximated by the sum of non-relativistic Maxwellians. When there are strong relativistic electron contributions, this approach cannot be used anymore as here relativistic corrections to the cross sections are needed, which are usually not taken into account in most plasma codes.

10 Plasma Modelling

10.1 Available Codes

There are several spectral codes available that can be used to model hot plasmas. Without being complete, we mention here briefly the following codes. For a more extensive review, see Raymond (2005).

First, there are codes which calculate predominantly spectra in Collisional Ionisation Equilibrium:

1. CHIANTI[1] (named after the wine region near Firenze, Landi et al. 2006) evolved from early work by Landini and Monsignori Fossi (1970); originally, it focused on the calculation of Solar spectra and hence EUV lines are well represented in this code.
2. APEC[2] (Astrophysical Plasma Emission Code (Smith et al. 2001) evolved from the original Raymond and Smith (1977) code; originally, the main goal was to calculate the emissivity and cooling curve properly. This model is available in the popular fitting package XSPEC[3] (Arnaud 1996).
3. Mekal (after MEwe, KAastra & Liedahl; see Mewe et al. 1995) evolved from early work by Mewe and collaborators starting from 1970 onwards. The Mekal model is also incorporated in XSPEC.

The Mekal model is now part of the SPEX package,[4] (Kaastra et al. 1996b). SPEX is a spectral fitting package developed at SRON Utrecht based on the original CIE models of Mewe and collaborators. It has grown significantly and also contains NEI and a few simple PIE models; more models are under development. The package offers the opportunity to plot or tabulate also various physical parameters of interest of the model, apart from the spectrum, and the atomic physics used in the Mekal model has been significantly updated in the SPEX version as compared to the older version available in XSPEC. A majority of all plots shown in this paper were produced using SPEX.

There are also models focusing upon photo-ionised plasmas. We mention here two codes (apart from SPEX):

1. XSTAR[5] evolved from the original work of Kallman and McCray (1982), focusing upon the X-ray band. It is relatively straightforward to use and there are options to use it with XSPEC.
2. Cloudy[6] (Ferland et al. 1998) evolved from work by Ferland and collaborators that originally was meant to model optical emission line regions. It gradually evolved to the X-ray band. It is somewhat more complicated to use, but has more options than XSTAR.

10.2 Important Parameters

In order to calculate X-ray spectra for collisional plasmas, there is a set of parameters that determine the spectral shape and flux. The two most important parameters are the electron

[1] http://www.arcetri.astro.it/science/chianti/chianti.html.

[2] http://cxc.harvard.edu/atomdb.

[3] http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec.

[4] http://www.sron.nl/spex.

[5] http://heasarc.gsfc.nasa.gov/docs/software/xstar/xstar.html.

[6] http://www.nublado.org.

temperature T_e and emission measure Y of the gas. The emission measure is sometimes not well or even incorrectly defined. We follow here the definition used in SPEX, namely $Y = \int n_e n_H dV$ where n_e and n_H are the electron and hydrogen density (whether ionised or not). Sometimes people use "n^2" in this expression, but it is not always obvious whether ion density, baryon density, electron density, total particle density or something else is meant. The next important parameters are the abundances of the elements; these do not only determine the line emission spectrum, but also affect the continuum emission (Sect. 5.1). Note that the abundances also affect the precise values of n_e/n_H and n_i/n_H. For a fully ionised plasma with proto-solar abundances (Lodders 2003) these ratios are 1.201 and 1.097, respectively (for the older Anders and Grevesse (1989) abundances these ratios are 1.209 and 1.099).

For ionising plasmas (NEI) also the parameter $U = \int n_e dt$ (see (28)) is important, as this describes the evolution of a shocked or recombining gas towards equilibrium.

When high spectral resolution observations are available, other parameters are important, like the ion temperature and turbulent velocity. These parameters determine the width of the spectral lines (see Sect. 5.2.2 and Sect. 6.3). Note that in non-equilibrium plasmas, the electron and ion temperatures are not necessarily the same. In an extreme case, the supernova remnant SN 1006, Vink et al. (2003) measured an oxygen ion temperature that is 350 times larger than the electron temperature of 1.5 keV. SPEX has also options to approximate the spectra when there are non-Maxwellian tails to the electron distribution (Sect. 9). These tails may be produced by non-thermal acceleration mechanisms (Petrosian et al. 2008—Chap. 9, this issue).

In general, for the low densities encountered in the diffuse ISM of galaxies, in clusters and in the WHIM, the absolute electron density is not important (except for the overall spectral normalisation through the emission measure Y). Only in the higher density regions of for example stars X-ray line ratios are affected by density effects, but this is outside the scope of the present paper.

However, if the density becomes extremely low, such as in the lowest density parts of the WHIM, the gas density matters, as in such situations photoionisation by the diffuse radiation field of galaxies or the cosmic background becomes important. But even in such cases, it is not the spectral shape but the ionisation (im)balance that is affected most. In those cases essentially the ratio of incoming radiation flux to gas density determines the photoionisation equilibrium.

10.3 Multi-Temperature Emission and Absorption

Several astrophysical sources cannot be characterised by a single temperature. There may be various reasons for this. It may be that the source has regions with different temperatures that cannot be spatially resolved due to the large distance to the source, the low angular resolution of the X-ray instrument, or superposition of spectra along the line of sight. Examples of these three conditions are stellar coronae (almost point sources as seen from Earth), integrated cluster of galaxies spectra, or spectra along the line of sight through the cooling core of a cluster, respectively. But also in other less obvious cases multi-temperature structure can be present. For example Kaastra et al. (2004) showed using deprojected spectra that at each position in several cool core clusters there is an intrinsic temperature distribution. This may be due to unresolved smaller regions in pressure equilibrium with different densities and hence different temperatures.

The most simple solution is to add as many temperature components as are needed to describe the observed spectrum, but there are potential dangers to this. The main reason is

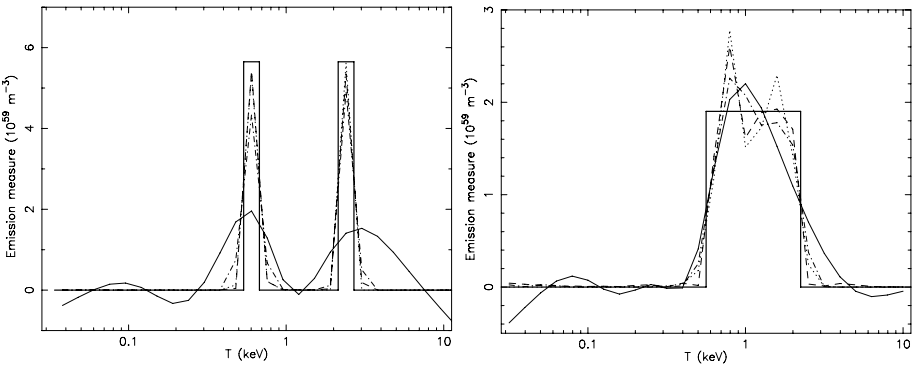


Fig. 16 Reconstructed DEM distribution for a simulated spectrum containing two equally strong isothermal components of 0.6 and 2.4 keV (*left panel*) or a uniform distribution between 0.6 and 2.4 keV (*right panel*). *Thick solid line*: the input model; *thin solid line*: regularisation method; *dot-dashed line*: polynomial method; *dashed line*: genetic algorithm; *dotted line*: two broadened Gaussian components. The corresponding best-fit spectra differ in χ^2 by less than 2 for each case. Obviously, small scale detail cannot be recovered but bulk properties such as the difference between a bimodal DEM distribution (*left panel*) or a flat distribution (*right panel*) can be recovered. From (Kaastra et al. 1996a). See that paper for more details

temperature sensitivity. Most spectral temperature indicators (line fluxes, Bremsstrahlung exponential cut-off) change significantly when the temperature is changed by a factor of two. However, for smaller temperature changes, the corresponding spectral changes are also minor. For example, a source with two components of equal emission measure Y and $kT = 0.99$ and 1.01 keV, respectively, is almost indistinguishable from a source with the emission measure $2Y$ and temperature 1.00 keV. Obviously, the most general case is a continuous differential emission measure distribution (DEM) $D(T) \equiv \mathrm{d}Y(T)/\mathrm{d}T$. Typically, in each logarithmic temperature range with a width of a factor of two, one can redistribute the differential emission measure $D(T)$ in such a way that the total emission measure as integrated over that interval as well as the emission-measure weighted temperature remain constant. The resulting spectra are almost indistinguishable. This point is illustrated in Fig. 16.

There are several methods to reconstruct $D(T)$. We refer here to Kaastra et al. (1996a) for more details. All these methods have been implemented in SPEX. They have been applied to stellar coronae or Solar spectra, but recently also to clusters of galaxies. See for example Kaastra et al. (2004). In the last case, an analytical power-law approximation $D(T) \sim T^{1/\alpha}$ has been used.

Also for photo-ionised plasmas a similar technique can be applied. Instead of the temperature T, one uses the ionisation parameter ξ (XSTAR users) or U (Cloudy users). In some cases (like Seyfert 2 galaxies with emission spectra) one recovers $D(\xi)$ from the observed emission spectra; in other cases (like Seyfert 1 galaxies with absorption spectra), a continuum model is applied to the observed spectrum, and $D(\xi)$ is recovered from the absorption features in the spectrum. Most of what has been said before on $D(T)$ also applies to this case, i.e. the resolution in ξ and the various methods to reconstruct the continuous distribution $D(\xi)$.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialised X-ray components in clusters of galaxies". SRON is supported financially by NWO, the Netherlands Organisation for Scientific Research.

References

E. Anders, N. Grevesse, Geochim. Cosmochim. Acta **53**, 197 (1989)
K.A. Arnaud, ASP Conf. Ser. **101**, 17 (1996)
M. Arnaud, R. Rothenflug, Astron. Astrophys. Suppl. Ser. **60**, 425 (1985)
A. Bhardwaj, R.F. Elsner, G.R. Gladstone, T.E. Cravens, C.M. Lisse, et al., Planet. Space Sci. **55**, 1135 (2006)
E. Churazov, M. Haehnelt, O. Kotov, R. Sunyaev, Mon. Not. R. Astron. Soc. **323**, 93 (2001)
G.J. Ferland, K.T. Korista, D.A. Verner et al., Publ. Astron. Soc. Pac. **110**, 761 (1998)
M.R. Gil'fanov, R.A. Syunyaev, E.M. Churazov, Sov. Astron. Lett. **13**, 3 (1987)
G. Herzberg, *Atomic Spectra and Atomic Structure*, 2nd edn. (Dover, New York, 1944)
J.S. Kaastra, R. Mewe, Astron. Astrophys. Suppl. Ser. **97**, 443 (1993)
J.S. Kaastra, R. Mewe, D.A. Liedahl, Astron. Astrophys. **314**, 547 (1996a)
J.S. Kaastra, R. Mewe, H. Nieuwenhuijzen, in *UV and X-ray Spectroscopy of Astrophysical and Laboratory Plasmas*, ed. by K. Yamashita, T. Watanabe (Univ. Ac. Press, Tokyo, 1996b), p. 411
J.S. Kaastra, T. Tamura, J.R. Peterson et al., Astron. Astrophys. **413**, 415 (2004)
T.R. Kallman, R. McCray, Astrophys. J. Suppl. Ser. **50**, 263 (1982)
W.J. Karzas, R. Latter, Astrophys. J. Suppl. Ser. **6**, 167 (1961)
K.D. Kuntz, S.L. Snowden, Astrophys. J. **543**, 195 (2000)
E. Landi, G. Del Zanna, P.R. Young et al., Astrophys. J. Suppl. Ser. **162**, 261 (2006)
M. Landini, B.C. Monsignori Fossi, Astron. Astrophys. **6**, 468 (1970)
K. Lodders, Astrophys. J. **591**, 1220 (2003)
W. Lotz, Z. Physik **216**, 441 (1968)
U. Maio, K. Dolag, B. Ciardi, L. Tornatore, Mon. Not. R. Astron. Soc. **379**, 963 (2007)
R. Mewe, Sol. Phys. **22**, 459 (1972)
R. Mewe, in *X-ray Spectroscopy in Astrophysics*. Lect. Notes in Phys., vol. 520 (1999), p. 109
R. Mewe, J.R. Lemen, G.H.J. van den Oord, Astron. Astrophys. Supp. **65**, 511 (1986)
R. Mewe, J.S. Kaastra, D.A. Liedahl, Legacy **6**, 16 (1995)
V. Petrosian, A. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9327-2
J.C. Raymond, AIP Conf. Proc. **774**, 15 (2005)
J.C. Raymond, B.W. Smith, Astrophys. J. Suppl. Ser. **35**, 419 (1977)
P. Richter, F. Paerels, J.S. Kaastra, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9325-4
M. Sako, S.M. Kahn, E. Behar et al., Astron. Astrophys. **365**, L168 (2001)
M.J. Seaton, Mon. Not. R. Astron. Soc. **119**, 81 (1959)
R.K. Smith, N.S. Brickhouse, D.A. Liedahl, J.C. Raymond, Astrophys. J. **556**, L91 (2001)
R.S. Sutherland, M.A. Dopita, Astrophys. J. Suppl. Ser. **88**, 253 (1993)
J. Vink, J.M. Laming, M.F. Gu, A. Rasmussen, J.S. Kaastra, Astrophys. J. **587**, L31 (2003)
J. Wilms, A. Allen, R. McCray, Astrophys. J. **542**, 914 (2000)
Y. Yao, Q.D. Wang, Astrophys. J. **641**, 930 (2006)

Chapter 10
Nonthermal Radiation Mechanisms

Vahé Petrosian · Andrei Bykov · Yoel Rephaeli

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9327-2

Abstract In this paper we review the possible radiation mechanisms for the observed non-thermal emission in clusters of galaxies, with a primary focus on the radio and hard X-ray emission. We show that the difficulty with the non-thermal, non-relativistic Bremsstrahlung model for the hard X-ray emission, first pointed out by Petrosian (Astrophys. J. **557**, 560, 2001) using a cold target approximation, is somewhat alleviated when one treats the problem more exactly by including the fact that the background plasma particle energies are on average a factor of 10 below the energy of the non-thermal particles. This increases the lifetime of the non-thermal particles, and as a result decreases the extreme energy requirement, but at most by a factor of three. We then review the synchrotron and so-called inverse Compton emission by relativistic electrons, which when compared with observations can constrain the value of the magnetic field and energy of relativistic electrons. This model requires a low value of the magnetic field which is far from the equipartition value. We briefly review the possibilities of gamma-ray emission and prospects for *GLAST* observations. We also present a toy model of the non-thermal electron spectra that are produced by the acceleration mechanisms discussed in an accompanying paper Petrosian and Bykov (Space Sci. Rev., 2008, this issue, Chap. 11).

Keywords Radiation mechanisms: non-thermal · Magnetic fields · Galaxies: clusters: general · X-rays: galaxies: clusters

V. Petrosian (✉)
Department of Applied Physics, Stanford University, Stanford, CA 94305, USA
e-mail: vahep@stanford.edu

V. Petrosian
Kavli Institute of Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, USA

A. Bykov
A.F. Ioffe Institute for Physics and Technology, 194021 St. Petersburg, Russia

Y. Rephaeli
School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel

1 Introduction

The observed signatures of the non-thermal (NT) activity in the intra-cluster medium (ICM) were described in details by Rephaeli et al. (2008—Chap. 5, this issue) and Ferrari et al. (2008—Chap. 6, this issue). Here we give a brief summary. The first and least controversial radiative signature comes from radio observations. Synchrotron emission by a population of relativistic electrons is the only possible model for the production of this radiation. In the case of the Coma cluster, the radio spectrum may be represented by a broken power law (Rephaeli 1979), or a power law with a rapid steepening (Thierbach et al. 2003), or with an exponential cutoff (Schlickeiser et al. 1987), implying the presence of electrons with similar spectra. Unfortunately, from radio observations alone one cannot determine the energy of the electrons or the strength of the magnetic field. Additional observations or assumptions are required. Equipartition or minimum total (relativistic particles plus fields) energy arguments imply a population of relativistic electrons with Lorentz factor $\gamma \sim 10^4$ and magnetic field strength of $B \sim \mu$G, in rough agreement with the Faraday rotation measurements (e.g., Kim et al. 1990). Rephaeli (1979) and Schlickeiser et al. (1987) also pointed out that the electrons responsible for the radio emission, should also produce a spectrum of hard X-ray (HXR) photons (similar to that observed in the radio band), via inverse Compton (IC) scattering of the Cosmic Microwave Background (CMB) photons. This emission is estimated to be the dominant emission component around 50 keV. Detection of HXR radiation could break the degeneracy and allow determination of the magnetic field and the energy of the radiating electrons. In fact, because the energy density of the CMB radiation (temperature T_0) $u_{\rm CMB} = 4 \times 10^{-13} (T_0/2.8\,{\rm K})^4$ erg cm^{-3} is larger than the magnetic energy density $u_B = 3 \times 10^{-14} (B/\mu{\rm G})^2$ erg cm^{-3}, one expects a higher flux of HXR than radio radiation.

As already described in the above mentioned papers by Rephaeli et al. and Ferrari et al., recently there has been growing evidence for this and other signatures of the NT activity. Excess HXR and extreme ultraviolet (EUV) radiation are observed at the high and low ends of the usual soft X-ray (SXR) thermal Bremsstrahlung (TB) radiation. Figure 1 shows all the flux $\nu F(\nu)$ (or equivalently the energy density $\nu u(\nu) = 4\nu F(\nu)/c$) of the above mentioned and other radiation for the Coma cluster. However, for the excess radiation not only the exact mechanisms are controversial but even their NT nature is questioned. The observed spectra of the excess radiation often can be fit by thermal spectra with higher and lower temperatures than that needed for the SXR observations with almost the same confidence as with a NT power law. The most natural NT process for these excesses (specially for HXRs) is the IC scattering of the CMB photons. However, the relatively high observed HXR fluxes require a large number of relativistic electrons, and consequently a relatively low magnetic field for a given observed radio flux. For Coma, this requires the (volume averaged) magnetic field to be $\bar{B} \sim 0.1$–0.3 μG, while equipartition gives $\bar{B} \sim 0.4$ μG and Faraday rotation measurements give the (average line-of-sight) field of $\bar{B} \sim 3$ μG (Giovannini et al. 1993; Kim et al. 1990; Clarke et al. 2001; Clarke 2003). In general the Faraday rotation measurements of most clusters give $B > 1$ μG; see e.g. Govoni et al. (2003). Because of this apparent difficulty, various authors (see, e.g. Enßlin et al. 1999; Blasi 2000) suggested that the HXR radiation is due to non-thermal Bremsstrahlung by a second population of NT electrons with a power law distribution in the 10 to 100 keV range. In what follows we examine the merits and shortcomings of the mechanisms proposed to interpret these observations. We first consider the EUV observations briefly and then address the thermal and NT (IC and non-thermal Bremsstrahlung) radiation model for the HXR observations.

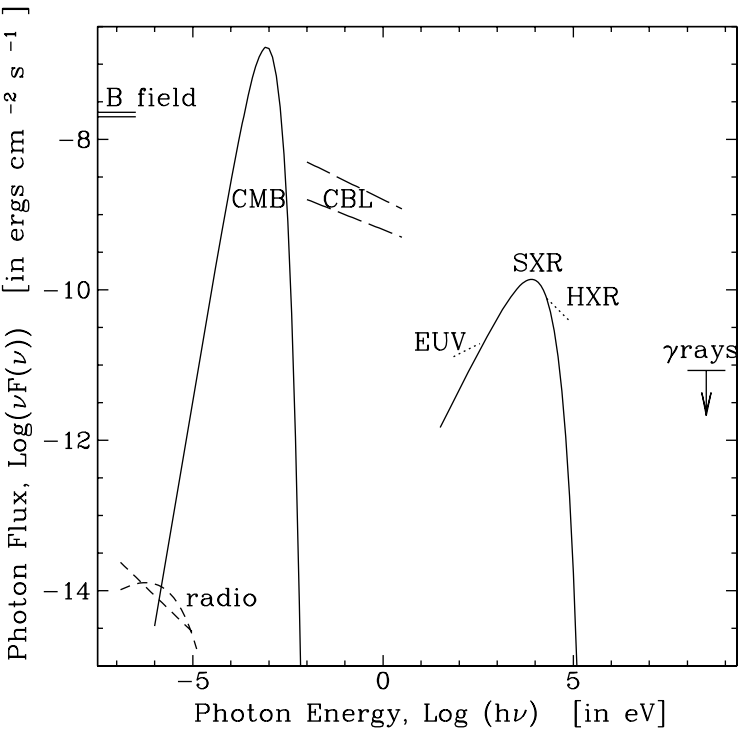


Fig. 1 The flux of all observed electromagnetic radiation for the Coma cluster including cosmic microwave background (CMB), cosmic background light (CBL) and static magnetic field (obtained from their energy density $u(\nu)$ as $\nu F(\nu) = (c/4) \times \nu u(\nu)$). The spectra shown for the EUV and HXR range are schematic and the upper limit in gamma ray range is from *EGRET* (Sreekumar et al. 1996). From Petrosian (2003)

2 EUV Emission

The EUV excess in the 0.07 to 0.14 keV range was first detected by the Extreme Ultraviolet Explorer from Coma (Lieu et al. 1996) and some other clusters. There are claimed detections of similar excess emissions in the 0.1 to 0.4 keV band by *Rosat*, *BeppoSAX* and *XMM-Newton*. The observational problems related to the EUV and soft excesses are discussed by Durret et al. (2008—Chap. 4, this issue). Initially, these excesses were attributed to thermal emission by a cooler ($kT \sim 2$ keV) component, but there are several theoretical arguments against this possibility, most notable is that the expected line emission is not observed. The alternative model is the IC scattering by CMB photons, which, in principle, can be easily fitted over the small range of observations. However, this will require a population of lower energy ($\gamma \sim 10^3$) electrons, indicating that the power law distribution required for production of radio radiation must be extended to lower energies with a power law index $p \sim 3$. This of course will mean an order of magnitude more energy in electrons and it makes equipartition less likely (see also the discussion at the end of Sect. 3.2.2).

In summary, some of the observations of the EUV emission are widely questioned and their theoretical modelling is quite problematic.

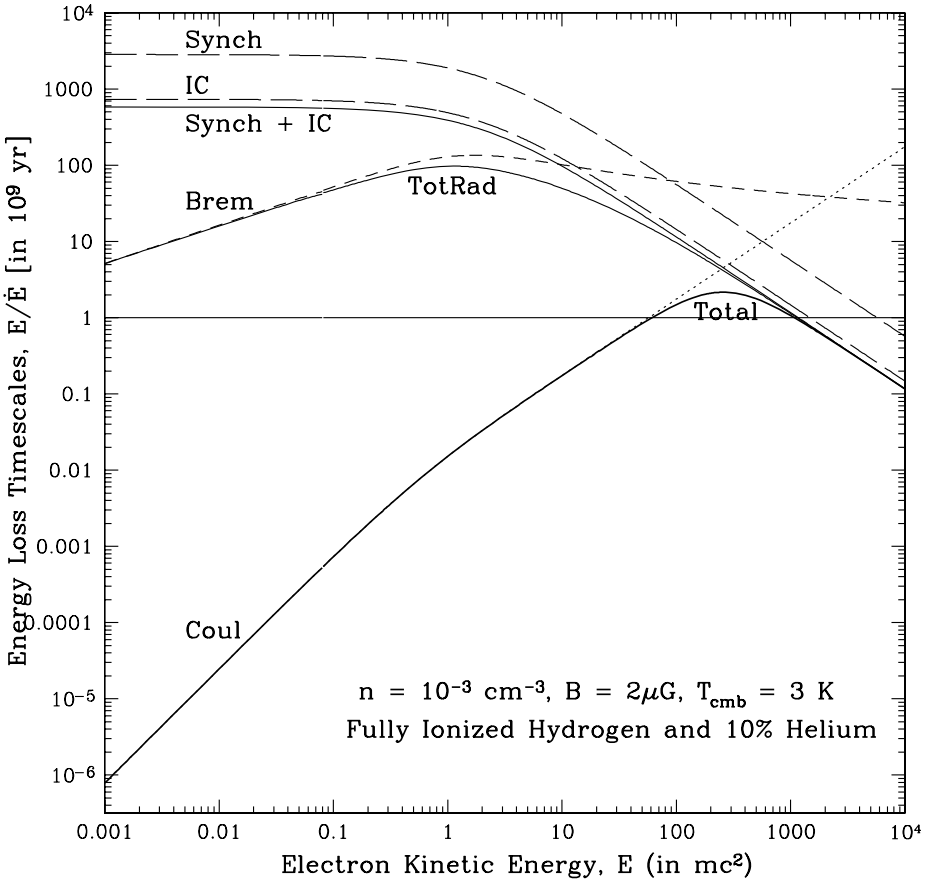


Fig. 2 Radiative, cold target Coulomb collision loss and other relevant timescales as a function of energy for the specified ICM conditions. The *solid lines* from top to bottom give the IC plus synchrotron, total radiative and total loss timescales. From Petrosian (2001)

3 Hard X-ray Emission

The two possibilities here are Bremsstrahlung (thermal or non-thermal) emission by a non-relativistic electron population and IC emission by extreme relativistic electrons which are also responsible for the observed radio emission. There are difficulties in both cases but as we show below the IC is the most likely interpretation. Figure 2 gives the timescales for the relevant processes for typical ICM conditions.

3.1 Bremsstrahlung Emission

Bremsstrahlung radiation in the HXR (20 to 100 keV) range requires electrons of comparable or somewhat larger energies. As pointed out by Petrosian (2001, P01 for short), and as evident from Fig. 2, for such nonrelativistic electrons NT Bremsstrahlung losses are negligible compared to elastic Coulomb losses. A thermal Bremsstrahlung interpretation of this emission requires temperatures well above the virial value (see also below). Very generally, for a particle of energy E interacting with background electrons and

protons of much lower energy (cold target), the energy yield in Bremsstrahlung photons $Y_{\rm brem} \equiv \dot{E}_{\rm brem}/\dot{E}_{\rm Coul} = (4\alpha \ln\Lambda/3\pi)(E/m_{\rm e}c^2) \sim 3 \times 10^{-6}(E/25\ {\rm keV})$. Here α is the fine structure constant and $\ln\Lambda$, the Coulomb logarithm, is ~ 40 for ICM conditions (for details see Petrosian 1973). Note that this result is independent of the spectrum of emitting electrons which could be a Maxwellian of higher temperature ($T_{\rm hot} \sim 30$ keV) or a power law of mean energy $\bar{E} \gg kT$ of the background particles.[1] As pointed out in P01, for continuous production of a HXR luminosity of $L_{\rm HXR}$ ($\sim 4 \times 10^{43}$ erg s^{-1} estimated for Coma, see Table 1), a power of $L_{\rm HXR}/Y_{\rm Brem}$ ($\sim 10^{49}$ erg s^{-1} for Coma) must be continuously fed into the ICM. This will increase the ICM temperature to $T \sim 10^8$ K after 6×10^7 yr, or to 10^{10} K in a Hubble time. An obvious conclusion here is that the HXR Bremsstrahlung emission phase must be very short lived. It should also be noted that such a hot gas or such high energy electrons cannot be confined in the ICM by gravity and will escape it in a crossing time of $\sim 3 \times 10^6$ yr unless it is confined by the magnetic field or by scattering. From Fig. 2 we see that the Coulomb scattering time, which is comparable to the Coulomb loss time, is equal to the crossing time so that the escape time of the particles will be comparable to these. Therefore, for confinement for periods exceeding the loss time we need a shorter scattering mean free path or timescale. For example, for the scattering timescale shown in this figure the escape time will be larger than the total electron loss time at all energies.

These estimates of short timescales are based on energy losses of electrons in a cold plasma which is an excellent approximation for electron energies $E \gg {\rm k}T$ (or for $T_{\rm hot} \gg T$). As E nears kT the Coulomb energy loss rate increases and the Bremsstrahlung yield decreases. For $E/{\rm k}T > 4$ we estimate that this increase will be at most about a factor of 3 (see below).

An exact treatment of the particle spectra in the energy regime close to the thermal pool requires a solution of non-linear kinetic equations instead of the quasi-linear approach justified for high energy particles where they can be considered as test particles. Apart from Coulomb collisions, also collisionless relaxation processes could play a role in the shaping of supra-thermal particle spectra (i.e. in the energy regime just above the particle kinetic temperature). We have illustrated the effect of fast collisionless relaxation of ions in a post-shock flow in Figs. 1 and 2 of Bykov et al. (2008—Chap. 8, this issue). The relaxation problem has some common features with N-body simulations of cluster virialisation discussed in other papers of this volume where collisionless relaxation effects are important. The appropriate kinetic equations to describe both of the effects are non-linear integro-differential equations and thus by now some simplified approximation schemes were used.

In a recent paper Dogiel et al. (2007) using linear Fokker-Planck equations for test particles concluded that in spite of the short lifetime of the test particles the "particle distribution" lifetime is longer and a power law tail can be maintained by stochastic turbulent acceleration without requiring the energy input estimated above. In what follows we address this problem not with the test particles and cold plasma assumption, but by carrying out a calculation of lifetimes of NT high energy tails in the ICM plasma. The proposed approach is based on still linearised, but more realistic kinetic equations for Coulomb relaxation of an initial distribution of NT particles.

[1] As shown in Petrosian and Bykov (2008—Chap. 11, this issue), also from the acceleration point of view for this scenario, the thermal and NT cases cannot be easily distinguished from each other. The acceleration mechanism energises the plasma and modifies its distribution in a way that both heating and acceleration take place.

3.1.1 Hot Plasma Loss Times and Thermalisation

The energy loss rate or relaxation into a thermal distribution of high energy electrons in a magnetised plasma can be treated by the *Fokker-Planck* transport equation for the gyrophased average distribution along the length s of the field lines $F(t, E, \mu, s)$, where μ stands for the pitch angle cosine. Assuming an isotropic pitch angle distribution and a homogeneous source (or more realistically integrating over the whole volume of the region) the transport equation describing the pitch angle averaged spectrum, $N(E,t) \propto \int\int F \mathrm{d}s\mathrm{d}\mu$, of the particles can be written as (see Petrosian and Bykov 2008—Chap. 11, this issue for more details):

$$\frac{\partial N}{\partial t} = \frac{\partial^2}{\partial E^2}[D_{\mathrm{Coul}}(E)N] + \frac{\partial}{\partial E}[\dot{E}_{\mathrm{Coul}}N], \tag{1}$$

where $D_{\mathrm{Coul}}(E)$ and $\dot{E}_{\mathrm{Coul}}(E)$ describe the energy diffusion coefficient and energy loss (+) or gain (−) rate due to Coulomb collisions. We have ignored Bremsstrahlung and IC and synchrotron losses, which are negligible compared to Coulomb losses at low energies and for typical ICM conditions.

As mentioned above, the previous analysis was based on an energy loss rate due to Coulomb collisions with a "cold" ambient plasma (target electrons having zero velocity):

$$\dot{E}^{\mathrm{cold}}_{\mathrm{Coul}} = m_{\mathrm{e}}\mathrm{c}^2/(\tau_{\mathrm{Coul}}\beta), \quad \text{where } \tau_{\mathrm{Coul}} = (4\pi \mathrm{r}_0^2 \ln\Lambda \mathrm{c}n)^{-1}, \tag{2}$$

$v = \mathrm{c}\beta$ is the particle velocity and $\mathrm{r}_0 = \mathrm{e}^2/(m_{\mathrm{e}}\mathrm{c}^2) = 2.8 \times 10^{-13}$ cm is the classical electron radius. For ICM density of $n = 10^{-3}$ cm^{-3}, and a Coulomb logarithm $\ln\Lambda = 40$, $\tau_{\mathrm{Coul}} = 2.7 \times 10^7$ yr. Note that the diffusion coefficient is zero for a cold target so that $\frac{\partial N}{\partial t} = \tau^{-1}_{\mathrm{Coul}} \frac{\partial}{\partial E}[N/\beta]$ from which one can readily get the results summarised above. As stated above this form of the loss rate is a good approximation when the NT electron velocity $v \gg v_{\mathrm{th}}$, where $v_{\mathrm{th}} = \sqrt{2\mathrm{k}T/m_{\mathrm{e}}}$ is the thermal velocity of the background electrons. This approximation becomes worse as $v \to v_{\mathrm{th}}$ and breaks down completely for $v < v_{\mathrm{th}}$, in which case the electron may gain energy rather than lose energy as is always the case in the cold-target scenario. A more general treatment of Coulomb loss is therefore desired. Petrosian and East (2008, PE07 herafter) describe such a treatment. We summarise their results below. For details the reader is referred to that paper and the references cited therein.

Let us first consider the *energy loss rate*. This is obtained from the rate of exchange of energy between two electrons with energies E and E' which we write as

$$\langle \Delta E \rangle / \Delta t = m_{\mathrm{e}}\mathrm{c}^2 G(E, E')/\tau_{\mathrm{Coul}}. \tag{3}$$

Here $G(E, E')$ is an antisymmetric function of the two variables such that the higher energy electrons loses and the lower one gains energy (see e.g. Nayakshin and Melia 1998). From their (24)–(26) we can write

$$G(E,E') = \begin{cases} -\beta'^{-1}, & \text{if } E' > E,\ E \ll m_{\mathrm{e}}\mathrm{c}^2; \\ \beta^{-1}, & \text{if } E' < E,\ E' \ll m_{\mathrm{e}}\mathrm{c}^2; \\ m_{\mathrm{e}}\mathrm{c}^2(E'^{-1} - E^{-1}), & \text{if } E, E' \gg m_{\mathrm{e}}\mathrm{c}^2. \end{cases} \tag{4}$$

The general Coulomb loss term is obtained by integrating over the particle distribution:

$$\dot{E}^{\mathrm{gen}}_{\mathrm{Coul}}(E,t) = \frac{m_{\mathrm{e}}\mathrm{c}^2}{\tau_{\mathrm{Coul}}} \int_0^\infty G(E,E')N(E',t)\mathrm{d}E'. \tag{5}$$

For non-relativistic particles this reduces to

$$\dot{E}^{\rm gen}_{\rm Coul}(E,t) = \dot{E}^{\rm cold}_{\rm Coul}\left(\int_0^E N(E',t){\rm d}E' - \int_E^\infty (\beta/\beta')N(E',t){\rm d}E'\right). \tag{6}$$

Similarly, the *Coulomb diffusion coefficient* can be obtained from

$$\frac{\langle(\Delta E)^2\rangle}{\Delta t} = \frac{(m_{\rm e}{\rm c}^2)^2 H(E,E')}{\tau_{\rm Coul}} \tag{7}$$

as $D^{\rm gen}_{\rm Coul}(E,t) = (m_{\rm e}^2{\rm c}^4/\tau_{\rm Coul})\int_0^\infty H(E,E')N(E',t){\rm d}E'$. From (35) and (36) of Nayakshin and Melia (1998)[2] we get

$$H(E,E') = \begin{cases} \beta^2/(3\beta'), & \text{if } E' > E, E \ll m_{\rm e}{\rm c}^2; \\ \beta'^2/(3\beta), & \text{if } E' < E, E' \ll m_{\rm e}{\rm c}^2; \\ 1/2, & \text{if } E, E' \gg m_{\rm e}{\rm c}^2. \end{cases} \tag{8}$$

Again, for nonrelativistic energies the Coulomb diffusion coefficient becomes

$$D^{\rm gen}_{\rm Coul}(E,t) = \frac{\dot{E}^{\rm cold}_{\rm Coul} m_{\rm e}{\rm c}^2\beta^2}{3}\left(\int_0^E (\beta'/\beta)^2 N(E',t){\rm d}E' + \int_E^\infty (\beta/\beta')N(E',t){\rm d}E'\right). \tag{9}$$

Thus, the determination of the distribution $N(E,t)$ involves solution of the combined integro-differential equations (1), (6) and (9), which can be solved iteratively. However, in many cases these equations can be simplified considerably. For example, if the bulk of the particles have a Maxwellian distribution

$$N(E') = n(2/\sqrt{\pi})({\rm k}T/m_{\rm e}{\rm c}^2)^{-3/2}E'^{1/2}{\rm e}^{-E'/{\rm k}T}, \tag{10}$$

with ${\rm k}T \ll m_{\rm e}{\rm c}^2$, then integrating over this energy distribution, and after some algebra, the net energy loss (gain) rate and the diffusion coefficient can be written as (see references in PE07):

$$\dot{E}^{\rm hot}_{\rm Coul} = \dot{E}^{\rm cold}_{\rm Coul}\left[{\rm erf}(\sqrt{x}) - 4\sqrt{\frac{x}{\pi}}e^{-x}\right], \tag{11}$$

and

$$D_{\rm Coul}(E) = \dot{E}^{\rm cold}_{\rm Coul}\left(\frac{{\rm k}T}{m_{\rm e}{\rm c}^2}\right)\left[{\rm erf}(\sqrt{x}) - 2\sqrt{\frac{x}{\pi}}e^{-x}\right], \quad \text{with } x \equiv \frac{E}{{\rm k}T}, \tag{12}$$

where ${\rm erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}{\rm d}t$ is the error function.

The numerical results presented below are based on another commonly used form of the transport equation in the code developed by Park and Petrosian (1995, 1996), where the diffusion term in (1) is written as $\frac{\partial}{\partial E}[D(E)\frac{\partial}{\partial E}N(E)]$. This requires modification of the loss term to

$$\dot{E}^{\rm eff}_{\rm Coul} = \dot{E}^{\rm hot}_{\rm Coul} + \frac{{\rm d}D_{\rm Coul}}{{\rm d}E} = \dot{E}^{\rm hot}_{\rm Coul}\left[1 - \frac{1}{x}\frac{1}{\gamma(\gamma+1)}\right], \tag{13}$$

[2]Note that the first term in their (35) should have a minus sign and that the whole quantity is too large by a factor of 2; see also Blasi (2000) for other typos.

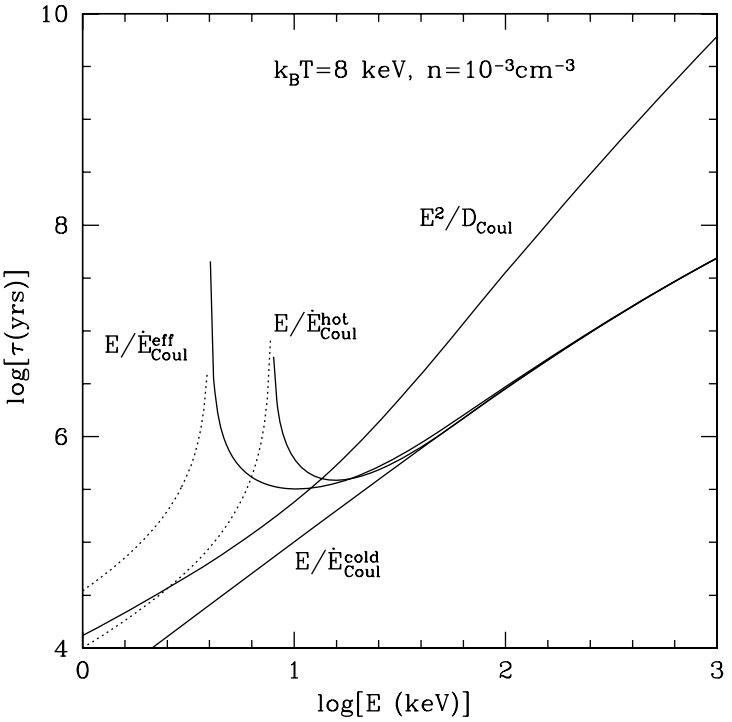


Fig. 3 Various timescales for Coulomb collisions for hot and cold plasma with typical ICM parameters from (2), (11), (12) and (13). Note that the as we approach the energy $E \rightarrow \mathrm{k}T$ the loss time increases and eventually becomes negative (or gain time) at low energies. The portions below the spike show the absolute value of these timescales. From Petrosian and East (2008)

where we have used (11) and (12). Figure 3 shows the loss and diffusion times,

$$\tau_{\mathrm{Coul}} = E/\dot{E}_{\mathrm{Coul}} \tag{14}$$

and

$$\tau_{\mathrm{diff}} = E^2/D_{\mathrm{Coul}}(E), \tag{15}$$

based on the above equations.

As a test of this algorithm PE07 show that an initially narrow distribution of particles (say, Gaussian with mean energy E_0 and width $\Delta E \ll E_0$) subject only to Coulomb collisions approaches a Maxwellian with $\mathrm{k}T = 2E_0/3$ with a constant total number of particles within several times the theoretically expected thermalisation time, which is related to our τ_{Coul} as (see Spitzer 1962 or Benz 2002)

$$\tau_{\mathrm{therm}} = 3.5\tau_{\mathrm{Coul}}(\mathrm{k}T/m_{\mathrm{e}}\mathrm{c}^2)^{1.5} = 2\tau_{\mathrm{Coul}}(E_0/m_{\mathrm{e}}\mathrm{c}^2)^{1.5}. \tag{16}$$

Using the above equations one can determine the thermalisation or the energy loss timescale of supra-thermal tails into a background thermal distribution. Figure 4 shows two examples of the evolution of an injected power-law distribution of electrons. The left panel shows this evolution assuming a constant temperature background plasma, which would be the case if either the energy of the injected electrons was negligible compared to that of

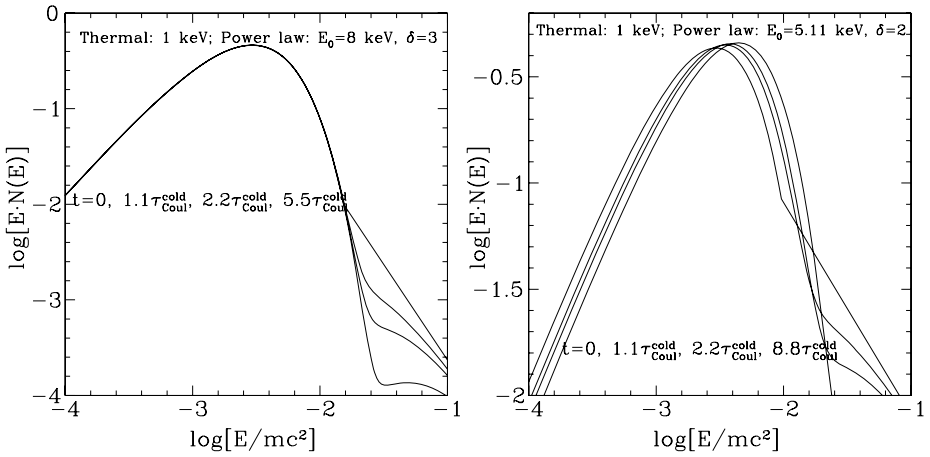


Fig. 4 Evolution of a NT power law tail (with isotropic angular distribution) of electrons subject to elastic Coulomb collisions with background thermal plasma electrons with initial temperature $kT = 1$ keV, showing gradual degradation of the power law NT tail (starting at energy E_0 with spectral index δ). *Left panel*: Here we assume that the energy input by the NT tail is negligible or carried away by some other means; i.e. the temperature of the plasma is forced to remain constant. *Right panel*: Here the energy of the NT particles remains in the system so that their thermalisation heats the plasma. In both cases the NT tails are reduced by a factor of ten within less than $100 \times \tau_{\rm therm}$ or < 3 times the cold target loss time at the injected energy. From Petrosian and East (2008)

the background electrons, or if the energy lost by the injected particles is lost by some other means. The right panel does not make these assumptions and allows the whole distribution to evolve. As is clearly seen in this figure, the energy lost by the NT particles heats the plasma. It is also evident that the NT tail is peeled away starting with low energies and progressing to higher ones. The NT tail becomes negligible within less than 100 times the thermalisation time of the background particles. These times are only about three times larger than the timescale one gets based on a cold target assumption for $E_0 = 20$ keV particles. The change in the electron lifetime agrees also with Fig. 3 of Dogiel et al. (2007). But our results do not support the other claims in their paper about the longer lifetime of the "distribution of particles" (for more details see PE07).

In summary the above results show that the conclusions based on the cold plasma approximation are good order of magnitude estimates and that using the more realistic hot plasma relations changes these estimates by factors of less than three. The upshot of this is that the required input energy will be lower by a similar factor and the time scale for heating will be longer by a similar factor compared to the estimates made in P01.

3.2 Emission from Relativistic Electrons

The radiative efficiency of relativistic electrons is much higher than that of nonrelativistic ones because of lower Coulomb losses (see Fig. 2). For ICM conditions electrons with GeV or higher energies lose energy via synchrotron and IC mechanisms. At such energies the Bremsstrahlung rate receives contributions from both electron-electron and electron-ion interactions and its loss rate becomes larger than Coulomb rate but remains below IC and synchrotron rates. Nevertheless, Bremsstrahlung emission could be the main source of gamma-ray radiation in the *GLAST* range ($>$ GeV). In this section we first compare the first two processes and explore whether the same population of electrons can be responsible for

both radio and HXR emissions. Then we consider Bremsstrahlung and other processes for the gamma-ray range.

3.2.1 Hard X-ray and Radio Emission

As stated in the introduction, relativistic electrons of similar energies ($\gamma > 10^3$) can be responsible for both the IC-HXR and synchrotron-radio emission. The IC and synchrotron fluxes depend on the photon (CMB in our case) and magnetic field energy densities and their spectra depend on the spectrum of the electrons (see e.g. Rybicki and Lightman 1979). For a power law distribution of relativistic electrons, $N(\gamma) = N_{\rm total}(p-1)\gamma^{-p}\gamma_{\rm min}^{p-1}$, and for $\gamma > \gamma_{\rm min}$, the spectrum of both radiation components, from a source at redshift $z \equiv Z - 1$ and co-moving coordinate $r(Z)$, is given by

$$\nu F_i(\nu) = c r_0^2 N_{\rm total}\gamma_{\rm min} u_i A_i(p)(\nu/\nu_{{\rm cr},i})^{2-\alpha}/(4\pi d_{\rm L}^2(Z), \tag{17}$$

where $d_{\rm L}(Z) = ({\rm c}/{\rm H}_0)Zr(Z)$ is the luminosity distance and $\alpha = (p+1)/2$ is the photon number spectral index.[3] This also assumes that the relativistic electrons and the magnetic field have similar spatial distribution so that the spatial distribution of HXRs and radio emission are the same (see e.g. Rephaeli 1979). The clusters are unresolved at HXRs so at this time this is the most reasonable and simplest assumption. A larger HXR flux will be the result if the electron and B field distributions are widely separated (see e.g., Brunetti et al. 2001).

For synchrotron $\nu_{\rm cr,synch} = 3\gamma_{\rm min}^2\nu_{B_\perp}/2$, with $\nu_{B_\perp} = {\rm e}B_\perp/(2\pi m_{\rm e}{\rm c})$ and $u_{\rm synch} = B^2/(8\pi)$, and for IC, u_i is the soft photon energy density and $\nu_{\rm cr,IC} = \gamma_{\rm min}^2\langle h\nu\rangle$. For black body photons $u_{\rm IC} = (8\pi^5/15)({\rm k}T)^4/({\rm hc})^3$ and $\langle {\rm h}\nu\rangle = 2.8\,{\rm k}T$. A_i are some simple functions of the electron index p and are of the order of unity and are given in Rybicki and Lightman (1979). Given the cluster redshift we know the temperature of the CMB photons ($T = T_0 Z$) and that $\nu_{\rm cr,IC} \propto Z$ and $u_{\rm IC} \propto Z^4$ so that we have

$$\mathcal{R} = F_{\rm HXR}/F_{\rm radio} \propto Z^{2+\alpha}/B_\perp^\alpha. \tag{18}$$

Consequently, from the observed ratio of fluxes we can determine the strength of the magnetic field.

For Coma, this requires the *volume averaged* magnetic field to be $\bar{B} \sim 0.1$ μG, while equipartition gives $\bar{B} \sim 0.4$ μG and Faraday rotation measurements give the *average line of sight* field of $\bar{B}_{\rm l} \sim 3$ μG (Giovannini et al. 1993; Kim et al. 1990; Clarke et al. 2001; Clarke 2003). In general the Faraday rotation measurements of most clusters give $B > 1$ μG; see e.g. Govoni et al. 2003. However, there are several factors which may resolve this discrepancy. Firstly, the last value assumes a chaotic magnetic field with a scale of a few kpc which is not a directly measured quantity (see e.g., Carilli and Taylor 2002).[4] Secondly, the accuracy of these results has been questioned (Rudnick and Blundell 2003; but see Govoni and Feretti 2004 for an opposing point of view). Thirdly, as shown by Goldshmidt and Rephaeli (1993), and also pointed out by Brunetti et al. (2001), a strong gradient in the magnetic field

[3] These expressions are valid for spectral index $p > 3$ or $\alpha > 2$. For smaller indices an upper energy limit $\gamma_{\rm max}$ must also be specified and the above expressions must be modified by other factors which are omitted here for the sake of simplification.

[4] The average line of sight component of the magnetic field in a chaotic field of scale $\lambda_{\rm chaos}$ will be roughly $\lambda_{\rm chaos}/R$ times the mean value of the magnetic field, where R is the size of the region.

can reconcile the difference between the volume and line-of-sight averaged measurements. Finally, as pointed out by P01, this discrepancy can be alleviated by a more realistic electron spectral distribution (e.g. the spectrum with exponential cutoff suggested by Schlickeiser et al. (1987) and/or a non-isotropic pitch angle distribution). In addition, for a population of clusters observational selection effects come into play and may favour Faraday rotation detection in high B clusters which will have a weaker IC flux relative to synchrotron. The above discussion indicates that the Faraday rotation measurements are somewhat controversial and do not provide a solid evidence against the IC model.

We now give some of the details relating the various observables in the IC model. We assume some proportional relation (e.g. *equipartition*) between the energies of the magnetic field and non-thermal electrons

$$\mathcal{E}_{\rm e} = N_{\rm total}\frac{p-1}{p-2}\gamma_{\rm min} m_{\rm e}{\rm c}^2 = \zeta\frac{B^2}{8\pi}\frac{4\pi R^3}{3}, \tag{19}$$

where $R = \theta d_{\rm A}/2$ is the radius of the (assumed) spherical cluster with measured angular diameter θ and angular diameter distance $d_{\rm A}(Z) = ({\rm c}/{\rm H}_0)r(Z)/Z$, and equipartition with electrons is equivalent to $\zeta = 1$.

From the three equations (17), (18) and (19) we can determine the three unknowns B, $\mathcal{E}_{\rm e}$ (or $N_{\rm total}$) and $F_{\rm HXR}$ purely in terms of ζ, $\gamma_{\rm min}$, and the observed quantities (given in Table 1) z, θ and the radio flux $\nu F_{\rm radio}(\nu)$. The result is[5]

$$(B/\mu{\rm G})^{\alpha+2} = 0.20\zeta^{-1}\left(\frac{F(1.4{\rm GHz})}{\rm Jy}\right)\left(\frac{5'}{\theta}\right)^3\left(\frac{10^4}{\gamma_{\rm min}}\right)^{2\alpha-3}\frac{Z^{3-\alpha}}{r(Z)}, \tag{20}$$

$$N_{\rm total} = 2.3\times 10^{65}\frac{\alpha-3/2}{\alpha-1}\zeta\left(10^3\gamma_{\rm min}\right)\left(\frac{\theta}{5'}\right)^3\left(\frac{B}{\mu{\rm G}}\right)^2\left(\frac{r(Z)}{Z}\right)^3, \tag{21}$$

and

$$\epsilon F_{\rm HXR}(\epsilon) = 0.034\times F_0\left(\frac{N_{\rm total}}{10^{65}}\right)\left(\frac{10^4}{\gamma_{\rm min}}\right)^{2\alpha-4}\left(\frac{\epsilon}{5.9\,{\rm keV}}\right)^{2-\alpha}\left(\frac{Z}{r(Z)}\right)^2, \tag{22}$$

where we have defined $F_0 \equiv 10^{-11}$ erg cm^{-2} s^{-1}. Note also that in all these expressions one may use the radius of the cluster $R = 3.39$ Mpc $(\theta/5')r(Z)/Z$ instead of the angular radius θ.

Table 1 presents a list of clusters with observed HXR emission and some other promising candidates. To obtain numerical estimates of the above quantities in addition to the observables $F_{\rm radio}, \theta$ and redshift z we need the values of ζ and $\gamma_{\rm min}$. Very little is known about these two parameters and how they may vary from cluster to cluster. From the radio observations at the lowest frequency we can set an upper limit on $\gamma_{\rm min}$; for Coma e.g., assuming $B \sim \mu$G we get $\gamma_{\rm min} < 4\times 10^3$ ($\mu{\rm G}/B$). We also know that the cut off energy cannot be too low because for $p > 3$ most of the energy of electrons ($\mathcal{E}_{\rm e} = \int_{\gamma_{\rm min}}^{\infty}\gamma^2 N(\gamma)d\gamma$) resides in the low energy end of the spectrum. As evident from Fig. 2 electrons with $\gamma < 100$ will lose their energy primarily via Coulomb collisions and heat the ICM. Thus, extending the spectra below this energy will cause excessive heating. A conservative estimate will be $\gamma_{\rm min} \sim 10^3$.

[5]Here we have set the Hubble constant ${\rm H}_0 = 70$ km s^{-1} Mpc^{-1}, the CMB temperature ${\rm T}_0 = 2.8$ K, and the radio frequency $\nu = 1.4$ GHz. In general $B^{2+\alpha} \propto {\rm H}_0\nu^{\alpha-1}$, $N_{\rm total} \propto {\rm H}_0^{-3}$ and $F_{\rm HXR} \propto {\rm H}_0^2{\rm T}_0^{2+\alpha}$. We have also assumed an isotropic distribution of the electron pitch angles and set $B = B_\perp(4/\pi)$.

Table 1 Observed and estimated properties of clusters

Cluster	z	kT[a] (keV)	$F_{1.4\,\mathrm{GHz}}$[b] (mJy)	θ[c,b] (arcmin)	F_{SXR} (F_0)[f]	B[d] (μG)	F_{HXR}[e] (F_0)[f]
Coma	0.023	7.9	52	30	33	0.40	1.4 (1.6)
A 2256	0.058	7.5	400	12	5.1	1.1	1.8 (1.0)
1ES 0657−55.8	0.296	15.6	78	5	3.9	1.2	0.52 (0.5)
A 2219	0.226	12.4	81	8	2.4	0.86	1.0
A 2163	0.208	13.8	55	6	3.3	0.97	0.50 (1.1)
MCS J0717.5+3745	0.550	13	220	3	3.5	2.6	0.76
A 1914	0.171	10.7	50	4	1.8	1.3	0.22
A 2744	0.308	11.0	38	5	0.76	1.0	0.41

[a]From Allen and Fabian (1998), except 1ES 0657—55.8 data from Liang et al. (2000)

[b]Data for Coma from Kim et al. (1990); A 2256, A 2163, A 1914 and A 2744 from Giovannini et al. (1999); 1ES 0657—55.8 from Liang et al. (2000); A 2219 from Bacchi et al. (2003)

[c]Approximate largest angular extent

[d]Estimates based on equipartition

[e]Estimates assuming $\zeta\gamma_{\min} = 10^6$, with observed values in parentheses: Coma (Rephaeli et al. 1999; Rephaeli and Gruber 2002; Fusco-Femiano et al. 1999, 2003); Abell 2256 (Fusco-Femiano et al. 2000, 2004; Rephaeli and Gruber 2003); 1ES 0657—55.8 (Petrosian et al. 2006); Abell 2163 (Rephaeli et al. 2006) these last authors also give a volume averaged $B \sim 0.4 \pm 0.2$ μG

[f]$F_0 \equiv 10^{-11}$ erg cm^{-2} s^{-1}

Even less is known about ζ. The estimated values of the magnetic fields B for the simple case of $\alpha = 2$, equipartition (i.e. $\zeta = 1$) and low energy cut off $\gamma_{\min} = 10^3$ are given in the 7th column of Table 1. As expected these are of the order of a few μG; for significantly stronger field, the predicted HXR fluxes will be below what is detected (or even potentially detectable). For $\alpha = 2$ the magnetic field $B \propto (\zeta\gamma_{\min})^{-1/4}$ and $F_{\mathrm{HXR}} \propto (\zeta\gamma_{\min})^{1/2}$ so that for sub-μG fields and $F_{\mathrm{HXR}} \sim F_0$ we need $\zeta\gamma_{\min} \sim 10^6$. Assuming $\gamma_{\min} = 10^3$ and $\zeta = 10^3$ we have calculated the expected fluxes integrated in the range of 20–100 keV (which for $\alpha = 2$ is equal to $1.62 \times [20\,\mathrm{keV}\, F_{\mathrm{HXR}}(20\,\mathrm{keV})]$), shown on the last column of Table 1. The variation of this flux with redshift based on the observed parameters, θ and $F_{\mathrm{radio}}(\nu = 1.4\,\mathrm{GHz})$ of Coma and A 2256 are plotted in Fig. 5 for three values of $\alpha = 1.75$, 2.0 and 2.25 ($p = 2.5$, 3 and 3.5) and assuming a constant physical radius R which is a reasonable assumption. We also plot the same assuming a constant angular diameter. This could be the case due to observational selection bias if diffuse radio emission is seen mainly from sources with θ near the resolution of the telescopes. These are clearly uncertain procedures and can give only semi-quantitative measures. However, the fact that the few observed values (given in parenthesis) are close to the predicted values is encouraging.

In summary we have argued that the IC is the most likely process for production of HXR (and possibly EUV) excesses in clusters of galaxies. The high observed HXR fluxes however imply that the situation is very far from equipartition.

Our estimates indicate that $\zeta \sim 10^6/\gamma_{\min} > 10^3$. In addition to the caveats enumerated above we note that, as described by Petrosian and Bykov (2008—Chap. 11, this issue), the sources generating the magnetic fields and the high energy electrons may not be identical so that equipartition might not be what one would expect.

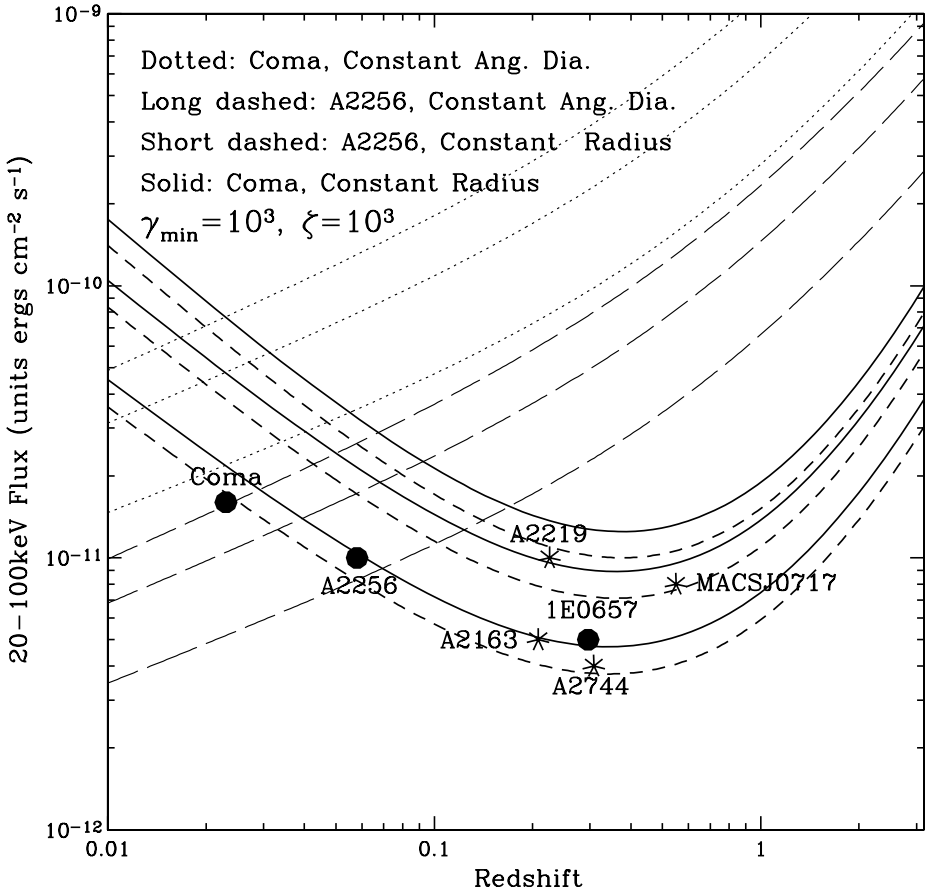


Fig. 5 Predicted variations of the HXR flux with redshift assuming a constant physical diameter (*solid* and *dashed lines*) or constant angular diameter (*dotted* and *long-dashed lines*), using the Coma cluster (*dotted* and *solid*) and A 2256 (*long or short dashed*) parameters assuming $\zeta = \gamma_{\min} = 10^3$. In each group, the IC photon spectral index $\alpha = 2.25, 2.0, 1.75$, from top to bottom. Filled circles are based on observations and stars based on the estimates given in Table 1. From Petrosian et al. (2006)

The spectrum of the required high energy electrons is best constrained by the radio observations. In Fig. 1 we showed two possible synchrotron spectra. The corresponding electron spectra are shown in Fig. 6 along with the Maxwellian distribution of thermal electrons. The low end of the NT spectra are constrained by requiring that their Coulomb loss rate be small so that there will not be an excessive heating of the background particles. These span the kind of NT electron spectrum that the acceleration model, discussed by Petrosian and Bykov 2008—Chap. 11, this issue, must produce.

3.2.2 Gamma-ray Emission

There have been several estimates of the expected gamma-ray flux, specifically from Coma (see e.g. Atoyan and Völk 2000; Bykov et al. 2000; Reimer et al. 2004; Blasi et al. 2007). The *EGRET* upper limit shown in Fig. 1 does not provide stringent constraints but *GLAST* with its much higher sensitivity can shed light on the processes described above. Several

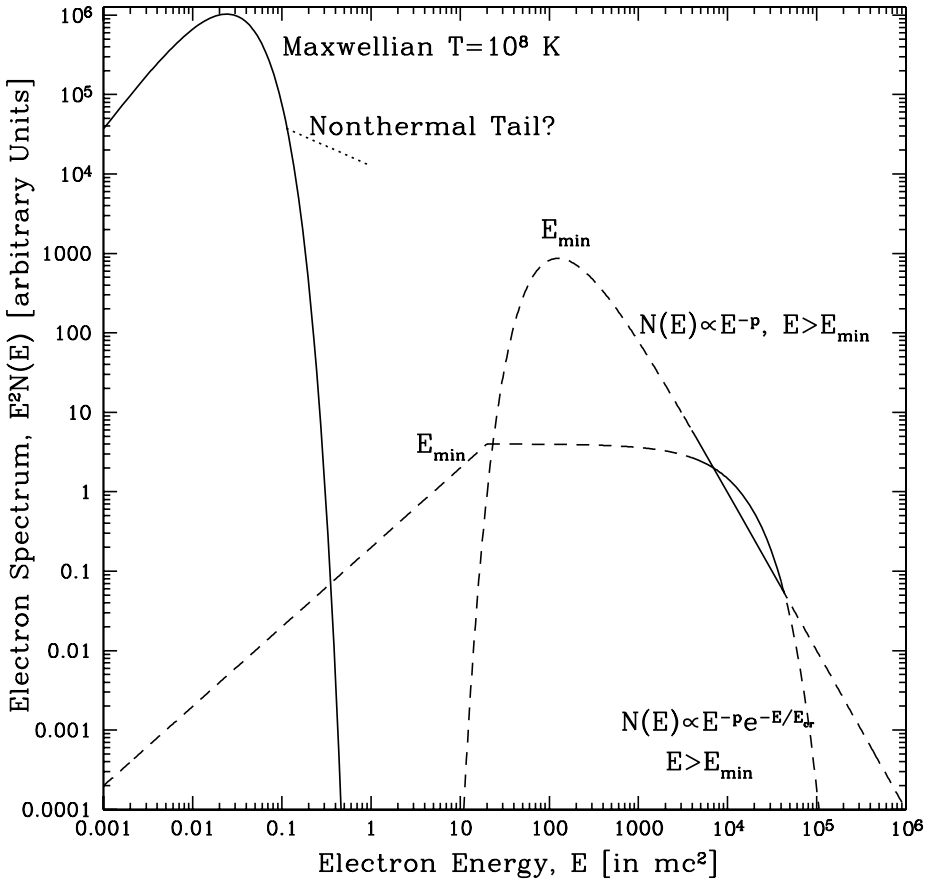


Fig. 6 The spectrum of electrons required for production of radio and HXR radiation based on the synchrotron and IC models. Values representing the Coma cluster are used. The low energy peak shows the distribution of thermal electrons along with a NT tail that would be required if the HXRs are produced by the non-thermal Bremsstrahlung mechanism. The low end of the high energy NT electrons are constrained to avoid excessive heating. The corresponding radio synchrotron spectra are shown in Fig. 1. From Petrosian (2003)

processes can produce gamma-ray radiation. Among electronic processes one expects IC and Bremsstrahlung radiation if electrons have sufficiently high energies. If the electron spectrum that accounts for radio and HXR emission can be extended to Lorentz factors of 3×10^5 (say with spectral index $p = 3$–4) then the IC scattering of CMB photons can produce 100 MeV radiation with a $\nu F(\nu)$ flux comparable to the HXR fluxes ($\sim 10^{-11}$–10^{-12} erg cm^{-2} s^{-1}) which can easily be detected by *GLAST* (the former actually disagrees with the *EGRET* upper limit shown in Fig. 1). The lifetime of such electrons is less than 10^7 years so that whatever the mechanism of their production (injection from galaxies or AGN, secondary electrons from proton-proton interactions, see the discussion by Petrosian and Bykov 2008—Chap. 11, this issue), these electrons must be produced throughout the cluster within times shorter than their lifetime. A second source could be electrons with $\gamma > 10^4$ scattering against the far infrared background (or cosmic background light) which will produce > 10 MeV photons but with a flux which is more than 100 times lower than

the HXR flux (assuming a cosmic background light to CMB energy density ratio of $< 10^{-2}$) which would be comparable to the *GLAST* one year threshold of $\sim 10^{-13}$ erg cm^{-2} s^{-1}. IC scattering of more abundant SXR photons by these electrons could produce > 100 MeV gamma-rays but the rate is suppressed by the Klein-Nishina effect. Only lower energy electrons $\gamma < 10^2$ will not suffer from this effect and could give rise to 100 MeV photons. As shown in Fig. 6 this is the lowest energy that a $p = 3$ spectrum can be extended to without causing excessive heating (see P01). These also may be the most abundant electrons because they have the longest lifetime (see Fig. 2). The expected gamma-ray flux will be greater than the *GLAST* threshold.

High energy electrons can also produce NT Bremsstrahlung photons with energies somewhat smaller than their energy. Thus, radio and HXR producing electrons will produce greater than 1 GeV photons. Because for each process the radiative loss rate roughly scales as $\dot{E} = E/\tau_{\rm loss}$, the ratio of the of gamma-ray to HXR fluxes will be comparable to the inverse of the loss times of Bremsstrahlung $\tau_{\rm brem}$ and IC $\tau_{\rm IC}$ shown in Fig. 2 ($\sim 10^{-2}$ for $\gamma \sim 10^4$), which is just above the *GLAST* threshold.

Finally hadronic processes by cosmic rays (mainly p–p scattering) can give rise to ~ 100 MeV and greater gamma-rays from the decay of π^0 produced in these scatterings. The $\pi^{\pm}$ decays give the secondary electrons which can contribute to the above radiation mechanisms. This is an attractive scenario because for cosmic ray protons the loss time is long and given an appropriate scattering agent, they can be confined in the ICM for a Hubble time (see, Berezinsky et al. 1997 and the discussion by Petrosian and Bykov 2008—Chap. 11, this issue). It is difficult to estimate the fluxes of this radiation because at the present time there are no observational constraints on the density of cosmic ray protons in the ICM. There is theoretical speculation that their energy density may be comparable to that of the thermal gas in which case the gamma-ray flux can be easily detected by *GLAST*. Reimer et al. (2004) estimate that *GLAST* can detect gamma-rays from the decay of π^0's if the cosmic ray proton energy is about 10% of the cluster thermal energy.

4 Summary

In this paper we address the processes that produce radio, EUV and HXR radiation in the ICM. First we consider NT Bremsstrahlung as the source of HXRs which, as shown earlier (P01), faces the difficulty of its low yield compared to Coulomb losses. We describe results from a more detailed analysis of the lifetimes of NT electron tails (or bumps) in a hot ICM than that presented in P01, where cold target loss rates were used. We find that the lifetimes of NT tails is increased by a factor of < 3 so that the above difficulty becomes less severe but production of HXRs via NT Bremsstrahlung remains problematical. Next we discuss the expected radiative signature of relativistic electrons and show that radio and HXR observations can be explained by synchrotron and IC scattering of CMB photons. But now one requires a low magnetic field which is far from equipartition with the electrons. Finally we give a rough estimate of the gamma-ray signature of the relativistic electrons and point out several possible scenarios in which the gamma-ray fluxes might exceed the *GLAST* threshold. Based on these results we present an average spectrum of electrons that is required and possible extensions of it into the low energy regime. Production of these spectra is discussed by Petrosian and Bykov (2008—Chap. 11, this issue).

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". A.M.B. acknowledges a support from RBRF grant 06-02-16844 and RAS Programs.

References

S.W. Allen, A.C. Fabian, Mon. Not. R. Astron. Soc. **297**, L57 (1998)
A.M. Atoyan, H.J. Völk, Astrophys. J. **535**, 45 (2000)
M. Bacchi, L. Feretti, G. Giovannini, F. Govoni, Astron. Astrophys. **400**, 465 (2003)
A. Benz, *Plasma Astrophysics. Kinetic Processes in Solar and Stellar Coronae*, 2nd edn. Astrophys. Sp. Sc. Lib., vol. 279 (Kluwer, Dordrecht, 2002)
V.S. Berezinsky, P. Blasi, V.S. Ptuskin, Astrophys. J. **487**, 529 (1997)
P. Blasi, Astrophys. J. **532**, L9 (2000)
P. Blasi, S. Gabici, G. Brunetti, Int. J. Mod. Phys. A **22**, 681 (2007)
G. Brunetti, G. Setti, L. Feretti, G. Giovannini, Mon. Not. R. Astron. Soc. **320**, 365 (2001)
A.M. Bykov, H. Bloemen, Yu.A. Uvarov, Astron. Astrophys. **362**, 886 (2000)
A.M. Bykov, F.B.S. Paerels, V. Petrosian, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9309-4
C.L. Carilli, G.B. Taylor, Annu. Rev. Astron. Astrophys. **40**, 319 (2002)
T.E. Clarke, P.P. Kronberg, H. Böhringer, Astrophys. J. **547**, L111 (2001)
T.E. Clarke, Astron. Soc. Pac. Conf. Ser. **301**, 185 (2003)
V.A. Dogiel, S. Colafrancesco, C.M. Ko et al., Astron. Astrophys. **461**, 433 (2007)
F. Durret, J.S. Kaastra, J. Nevalainen, T. Ohashi, N. Werner, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9313-8
T.A. Enßlin, R. Lieu, P.L. Biermann, Astron. Astrophys. **344**, 409 (1999)
C. Ferrari, F. Govoni, S. Schindler, A.M. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9311-x
R. Fusco-Femiano, D. Dal Fiume, L. Feretti et al., Astrophys. J. **513**, L21 (1999)
R. Fusco-Femiano, D. Dal Fiume, S. De Grandi et al., Astrophys. J. **534**, L7 (2000)
R. Fusco-Femiano, M. Orlandini, S. De Grandi et al., Astron. Astrophys. **398**, 441 (2003)
R. Fusco-Femiano, M. Orlandini, G. Brunetti et al., Astrophys. J. **602**, L73 (2004)
G. Giovannini, L. Feretti, T. Venturi, K.-T. Kim, P.P. Kronberg, Astrophys. J. **406**, 399 (1993)
G. Giovannini, M. Tordi, L. Feretti, New Astron. Rev. **4**, 141 (1999)
O. Goldshmidt, Y. Rephaeli, Astrophys. J. **411**, 518 (1993)
F. Govoni, L. Feretti, Int. J. Mod. Phys. D **13**, 1549 (2004)
F. Govoni, L. Feretti, M. Murgia et al., ASP Conf. Ser. **301**, 501 (2003)
K.T. Kim, P.P. Kronberg, P.E. Dewdney, T.L. Landecker, Astrophys. J. **355**, 29 (1990)
H. Liang, R.W. Hunstead, M. Birkinshaw, P. Andreani, Astrophys. J. **544**, 686 (2000)
R. Lieu, J.P.D. Mittaz, S. Bowyer et al., Astrophys. J. **458**, L5 (1996)
S. Nayakshin, F. Melia, Astrophys. J. Suppl. Ser. **114**, 269 (1998)
B.T. Park, V. Petrosian, Astrophys. J. **446**, 699 (1995)
B.T. Park, V. Petrosian, Astrophys. J. Suppl. Ser. **103**, 255 (1996)
V. Petrosian, Astrophys. J. **186**, 291 (1973)
V. Petrosian, Astrophys. J. **557**, 560 (2001)
V. Petrosian, ASP Conf. Ser. **301**, 337 (2003)
V. Petrosian, A.M. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9315-6
V. Petrosian, W. East, Astrophys. J. (2008), in press. arXiv:0802.0900
V. Petrosian, G. Madejski, K. Luli, Astrophys. J. **652**, 948 (2006)
A. Reimer, O. Reimer, R. Schlickeiser, A. Iyudin, Astron. Astrophys. **424**, 773 (2004)
Y. Rephaeli, Astrophys. J. **227**, 364 (1979)
Y. Rephaeli, D. Gruber, Astrophys. J. **579**, 587 (2002)
Y. Rephaeli, D. Gruber, Astrophys. J. **595**, 137 (2003)
Y. Rephaeli, D. Gruber, P. Blanco, Astrophys. J. **511**, L21 (1999)
Y. Rephaeli, D. Gruber, Y. Arieli, Astrophys. J. **649**, 673 (2006)
Y. Rephaeli, J. Nevalainen, T. Ohashi, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9314-7
L. Rudnick, K.M. Blundell, Astrophys. J. **588**, 143 (2003)
G.B. Rybicki, A.P. Lightman, *Radiative Processes in Astrophysics* (Wiley, New York, 1979)
R. Schlickeiser, A. Sievers, H. Thiemann, Astron. Astrophys. **182**, 21 (1987)
L. Spitzer, *Physics of Fully Ionized Gases*, 2nd edn. (Interscience, New York, 1962)
P. Sreekumar, D.L. Bertsch, B.L. Dingus et al., Astrophys. J. **464**, 628 (1996)
M. Thierbach, U. Klein, R. Wielebinski, Astron. Astrophys. **397**, 53 (2003)

Chapter 11
Particle Acceleration Mechanisms

V. Petrosian · A.M. Bykov

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9315-6

Abstract In this paper we review the possible mechanisms for production of non-thermal electrons which are responsible for the observed non-thermal radiation in clusters of galaxies. Our primary focus is on non-thermal Bremsstrahlung and inverse Compton scattering, that produce hard X-ray emission. We first give a brief review of acceleration mechanisms and point out that in most astrophysical situations, and in particular for the intracluster medium, shocks, turbulence and plasma waves play a crucial role. We also outline how the effects of the turbulence can be accounted for. Using a generic model for turbulence and acceleration, we then consider two scenarios for production of non-thermal radiation. The first is motivated by the possibility that hard X-ray emission is due to non-thermal Bremsstrahlung by nonrelativistic particles and attempts to produce non-thermal tails by accelerating the electrons from the background plasma with an initial Maxwellian distribution. For acceleration rates smaller than the Coulomb energy loss rate, the effect of energising the plasma is to primarily heat the plasma with little sign of a distinct non-thermal tail. Such tails are discernible only for acceleration rates comparable or larger than the Coulomb loss rate. However, these tails are accompanied by significant heating and they are present for a short time of $< 10^6$ years, which is also the time that the tail will be thermalised. A longer period of acceleration at such rates will result in a runaway situation with most particles being accelerated to very high energies. These more exact treatments confirm the difficulty with this model, first pointed out by Petrosian (Astrophys. J. 557:560, 2001). Such non-thermal tails, even if possible, can only explain the hard X-ray but not the radio emission which needs GeV or higher energy electrons. For these and for production of hard X-rays by the inverse Compton model, we need the second scenario where there is injection and subsequent acceleration of relativistic electrons. It is shown that a steady state situation, for example arising from secondary electrons produced from cosmic ray proton scattering by background protons, will most likely lead to flatter than required electron spectra or it requires a short escape

V. Petrosian (✉)
Department of Physics, Stanford University, Stanford, CA 94305, USA
e-mail: vahep@stanford.edu

A.M. Bykov
A.F. Ioffe Institute for Physics and Technology, St. Petersburg, 194021, Russia
e-mail: byk@astro.ioffe.ru

time of the electrons from the cluster. An episodic injection of relativistic electrons, presumably from galaxies or AGN, and/or episodic generation of turbulence and shocks by mergers can result in an electron spectrum consistent with observations but for only a short period of less than one billion years.

Keywords Intergalactic medium · Particle acceleration · Galaxies: clusters: general

1 Introduction

In this paper we attempt to constrain the acceleration models based on the observations described in the papers by Durret et al. (2008), Rephaeli et al. (2008) and Ferrari et al. (2008)—Chaps. 4–6, this issue, and the required spectrum of the accelerated electrons shown in Fig. 6 of Petrosian et al. (2008)—Chap. 10, this issue. However, before addressing these details we first compare various acceleration processes and stress the importance of plasma waves or turbulence (PWT) as an agent of scattering and acceleration, and then describe the basic scenario and equations for treatment of these processes. As pointed out below there is growing evidence that PWT plays an important role in acceleration of particles in general, and in clusters of galaxies in particular. The two most commonly used acceleration mechanisms are the following.

1.1 Electric Field Acceleration

Electric fields parallel to magnetic fields can accelerate charged particles and can arise as a result of magnetic field reconnection in a current sheet or other situations. For fields less than the so-called Dreicer field, defined as $E_{\rm D} = {\rm k}T/({\rm e}\lambda_{\rm Coul})$, where

$$\lambda_{\rm Coul} \sim 15\ {\rm kpc} \left(\frac{T}{10^8\ {\rm K}}\right)^2 \left(\frac{10^{-3}\ {\rm cm}^{-3}}{n}\right) \tag{1}$$

is the collision (electron-electron or proton-proton) mean free path,[1] the rate of acceleration is less than the rate of collision losses and only a small fraction of the particles can be accelerated into a non-thermal tail of energy $E < LeE_{\rm D}$. For the ICM $E_{\rm D} \sim 10^{-14}\ {\rm V\,cm}^{-1}$ and $L \sim 10^{24}$ cm so that sub-Dreicer fields can only accelerate particles up to 100's of keV, which is far below the 10's of GeV electrons required by observations. Super-Dreicer fields, which seem to be present in many simulations of reconnection (Drake 2006; Cassak et al. 2006; Zenitani and Hoshino 2005), accelerate particles at a rate that is faster than the collision or thermalisation time $\tau_{\rm therm}$. This can lead to a runaway and an unstable electron distribution which, as shown theoretically, by laboratory experiments and by the above mentioned simulations, most probably will give rise to PWT (Boris et al. 1970; Holman 1985).

In summary the electric fields arising as a result of reconnection cannot be the sole agent of acceleration in the ICM, because there are no large scale magnetically dominated cosmological flows, but it may locally produce an unstable particle momentum distribution which will produce PWT that can then accelerate particles.

[1] The proton-proton or ion-ion mean free path will be slightly smaller because of the larger value of the Coulomb logarithm $\ln \Lambda \sim 40$ in the ICM.

1.2 Fermi Acceleration

Nowadays this process has been divided into two kinds. In the original Fermi process particles of velocity v moving along magnetic field lines (strength B) with a pitch angle $\cos\mu$ undergo random scattering by moving agents with a velocity u. Because the head-on (energy gaining) collisions are more probable than trailing (energy losing) collisions, on average, the particles gain energy at a rate proportional to $(u/v)^2 D_{\mu\mu}$, where $D_{\mu\mu}$ is the pitch angle diffusion rate. This, known as a *second order Fermi process* is what we shall call stochastic acceleration. In general, the most likely agent for scattering is PWT. An alternative process is what is commonly referred to as a *first order Fermi process*, where the actual acceleration occurs when particles cross a shock or a region of converging flow. Upon crossing the shock the fractional gain of momentum $\delta p/p \propto u_{\rm sh}/v$. Ever since the 1970s, when several authors demonstrated that a very simple version of this process leads to a power law spectrum that agrees approximately with observations of the cosmic rays, shock acceleration is commonly invoked in space and astrophysical plasmas. However, this simple model, though very elegant, has some shortcomings specially when applied to electron acceleration in non-thermal radiating sources. Moreover, some of the features that make this scenario for acceleration of cosmic rays attractive are not present in most radiating sources where one needs efficient acceleration of electrons to relativistic energies from a low energy reservoir.

The original, test particle theory of diffusive shock acceleration (DSA), although very elegant and independent of geometry and other details (e.g. Blandford and Ostriker 1978) required several conditions such as injection of seed particles and of course turbulence. A great deal of work has gone into addressing these aspects of the problem and there has been a great deal of progress. It is clear that nonlinear effects (see e.g. Drury 1983; Blandford and Eichler 1987; Jones and Ellison 1991; Malkov and Drury 2001) and losses (specially for electrons) play an important role and modify the resultant spectra and efficiency of acceleration. Another important point is the source of the turbulence or the scattering agents. A common practice is to assume Bohm diffusion (see e.g. Ellison et al. 2005). In addition, second order acceleration effects could modify the particle spectra accelerated by shocks (see e.g. Schlickeiser et al. 1993; Bykov et al. 2000). Although there are indications that turbulence may be generated by the shocks and the accelerated particle upstream, many details (e.g. the nature and spectrum of the turbulence) need to be addressed more quantitatively. There has been progress on the understanding of generation of the magnetic field and turbulence on strong shocks (Bell and Lucek 2001; Amato and Blasi 2006; Vladimirov et al. 2006) as required in recent observations of supernova remnants (see e.g. Völk et al. 2005). There is also some evidence for these processes from observations of heliospheric shocks (see e.g. Kennel et al. 1986; Ellison et al. 1990). Basic features of particle acceleration by cosmological shocks were discussed by Bykov et al. (2008)—Chap. 7, this issue, so we will concentrate here on the stochastic acceleration perspective.

1.3 Stochastic Acceleration

The PWT needed for scattering can also accelerate particles stochastically with a rate D_{EE}/E^2, where D_{EE} is the energy diffusion coefficient, so that shocks may not be always necessary. In low beta plasmas, $\beta_{\rm p} = 2(v_{\rm s}/v_{\rm A})^2 < 1$ (here the Alfv́en velocity $v_{\rm A} = \sqrt{B^2/4\pi\rho}$, the sound velocity $v_{\rm s} = \sqrt{kT/m}$, $\rho = nm$ is the mass density and n is the number density of the gas) and for relativistic particles the PWT-particle interactions are dominated by Alfv́enic turbulence, in which case the rate of energy gain $D_{EE}/E^2 = (v_{\rm A}/v)^2 D_{\mu\mu} \ll D_{\mu\mu}$, so that the first order Fermi process is more efficient. However, at low energies and/or

in very strongly magnetised plasmas, where v_A can exceed c, the speed of light,[2] the acceleration rate may exceed the scattering rate (see Pryadko and Petrosian 1997), in which case low energy electrons are accelerated more efficiently by PWT than by shocks.[3]

Irrespective of which process dominates the particle acceleration, it is clear that PWT has a role in all of them. Thus, understanding of the production of PWT and its interaction with particles is extremely important. Moreover, turbulence is expected to be present in most astrophysical plasmas including the ICM and in and around merger or accretion shocks, because the ordinary and magnetic Reynolds numbers are large. Indeed turbulence may be the most efficient channel of energy dissipation. In recent years there has been a substantial progress in the understanding of MHD turbulence (Goldreich and Sridhar 1995, 1997; Lithwick and Goldreich 2003; Cho and Lazarian 2002, 2006). These provide new tools for a more quantitative investigation of turbulence and the role it plays in many astrophysical sources.

2 Turbulence and Stochastic Acceleration

2.1 Basic Scenario

The complete picture of stochastic acceleration by PWT is a complex and not yet fully understood or developed process. However, one might envision the following scenario.

Turbulence or plasma waves can be generated in the ICM on some macroscopic scale $L \sim 300$ kpc (some fraction of the cluster size or some multiple of galactic sizes) as a result of merger events or by accretion or merger shocks. That these kind of motions or flows with velocity comparable to or somewhat greater than the virial velocity $u_L \sim 1000\ \mathrm{km\,s^{-1}}$ will lead to PWT is very likely, because in the ICM the ordinary Reynolds number $R_e = u_L L/\nu \gg 1$. Here $\nu \sim v_{th}\lambda_{scat}/3$ is the viscosity, $v_{th} = \sqrt{kT/m} \sim u_L(T/10^8)^{1/2}$ and λ_{scat} is the mean free path length. The main uncertainty here is in the value of λ_{scat}. For Coulomb collisions $\lambda_{scat} \sim 15$ kpc (Eq. 1) and $R_e \sim 100$ is just barely large enough for generation of turbulence. However, in a recent paper Brunetti and Lazarian (2007) argue that in the presence of a magnetic field of $B \sim \mu$G, $v_A \sim 70(B/\mu\mathrm{G})(10^{-3}\ \mathrm{cm^{-3}}/n)^{1/2}\ \mathrm{km\,s^{-1}}$ is much smaller than v_{th} so that the turbulence will be super-Alfvénic, in which case the mean free path may be two orders of magnitude smaller[4] yielding $R_e \sim 10^4$. We know this also to be true from a phenomenological consideration. In a cluster the hot gas is confined by the gravitational field of the total (dark and 'visible') matter. Relativistic particles, on the other hand, can cross the cluster (radius R) on a timescale of $T_{cross} = 3 \times 10^6\ (R/\mathrm{Mpc})$ years and can escape the cluster (see Fig. 4 below), unless confined by a chaotic magnetic field or a scattering agent such as turbulence, with a mean free path $\lambda_{scat} \ll R$. If so, then the escape time $T_{esc} \sim T_{cross}(R/\lambda_{scat}) = T^2_{cross}/\tau_{scat}$. The curve marked with arrows in this figure shows the maximum value of the required τ_{scat} so that the escape time is longer than the energy loss time τ_{loss}. As is evident from this figure, for a GeV electron to be confined for a Hubble timescale, or $T_{esc} \sim 10^{10}$ years, we need $\tau_{scat} \sim 3 \times 10^4$ years or $\lambda_{scat} < 10$ kpc.

[2] Note that the Alfv́en group velocity $v_g = c\sqrt{v_A^2/(v_A^2 + c^2)}$ is always less than c.

[3] In practice, i.e. mathematically, there is little difference between the two mechanisms (Jones 1994), and the acceleration by turbulence and shocks can be combined (see below).

[4] Plasma instabilities, possibly induced by the relativistic particles, can be another agent of decreasing the effective particle mean free path (Schekochihin et al. 2005).

Some observations related to this are discussed by Petrosian et al. (2008)—Chap. 10, this issue; see also Vogt and Enßlin (2005). Numerous numerical simulations also agree with this general picture. There is evidence for large scale bulk flows in the simulations of merging clusters (e.g. Roettiger et al. 1996; Ricker and Sarazin 2001), and that these are converted into turbulence with energies that are a substantial fraction of the thermal energy of the clusters (e.g. Sunyaev et al. 2003; Dolag et al. 2005). For more details see Brunetti and Lazarian (2007).

Once the PWT is generated it can undergo two kind of interactions. The first is dissipationless cascade from wave vectors $k_{\rm min} \sim L^{-1}$ to smaller scales. The cascade is governed by the rates of *wave-wave interactions*. For example, in the case of weak turbulence, that can be considered as a superposition of weakly interacting wave packets, the three wave interactions can be represented as

$$\omega(\mathbf{k}_1) + \omega(\mathbf{k}_2) = \omega(\mathbf{k}_3) \quad \text{and} \quad \mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3, \tag{2}$$

where $\mathbf{k}$ is the wave vector, and the wave frequency, $\omega(\mathbf{k})$, is obtained from the *plasma dispersion relation*. One can interpret Eq. 2 as energy-momentum conservation laws for weakly coupled plasma waves in a close analogy to the optical waves. The interaction rates can be represented by the wave diffusion coefficient D_{ij} or the cascade time $\tau_{\rm cas} \sim k^2/D_{ij}$. The largest uncertainty is in the diffusion coefficient. Because of the nonlinear nature of the interactions this coefficient depends on the wave spectrum $W(\mathbf{k})$. As mentioned above there has been considerable progress in this area in the past two decades and there are some recipes how to calculate the diffusion coefficients.

The second is *damping of the PWT* by *wave-particle interaction* which terminates the dissipationless cascade, say at an outer scale $k_{\rm max}$ when the damping rate $\Gamma(k_{\rm max}) = \tau_{\rm cas}^{-1}(k_{\rm max})$. The range $k_{\rm min} < k < k_{\rm max}$ is called the inertial range. The damping rate can be obtained from the finite temperature dispersion relations (see below). The energy lost from PWT goes into heating the background plasma and/or accelerating particles into a non-thermal tail. These processes are described by the diffusion coefficients D_{EE} and $D_{\mu\mu}$ introduced above. These coefficients are obtained from consideration of the wave-particle interactions which are often dominated by resonant interactions, specially for low beta (magnetically dominated) plasma, such that

$$\omega(\mathbf{k}) - k\cos\theta v\mu = n\Omega/\gamma, \tag{3}$$

for waves propagating at an angle θ with respect to the large scale magnetic field, and a particle of velocity v, Lorentz factor γ, pitch angle $\cos\mu$ and gyrofrequency $\Omega = eB/mc$. Both cyclotron (the term in the right hand side of Eq. 3) and Cerenkov resonance (the second term in the left hand side) play important roles in the analysis (see for details e.g. Akhiezer et al. 1975). Here, when the harmonic number n (not to be confused with the density) is equal to zero, the process is referred to as the *transit time damping*. For gyroresonance damping by waves propagating parallel to the field lines ($\theta = 0$) $n = \pm 1$. For obliquely propagating waves, in principle one gets contributions from all harmonics $n = \pm 1, \pm 2, \ldots$, but for practical purposes most of the contribution comes from the lowest harmonics $n = \pm 1$ (see Pryadko and Petrosian 1998).

2.2 Dispersion Relations

It is clear from the above description that at the core of the evaluation of wave-wave or wave-particle interactions (and all the coefficients of the kinetic equations described below)

lies the plasma dispersion relation $\omega(\mathbf{k})$. It describes the characteristics of the waves that can be excited in the plasma, and the rates of wave-wave and wave-particle interactions.

In the MHD regime for a cold plasma

$$\omega = v_{\rm A} k \cos\theta \quad \text{and} \quad \omega = v_{\rm A} k \tag{4}$$

for the Alfvén and the fast (magneto-sonic) waves, respectively. Beyond the MHD regime a multiplicity of wave modes can be present and the dispersion relation is more complex and is obtained from the following expressions (see e.g. Sturrock 1994):

$$\tan^2\theta = \frac{-P(n_r^2 - R)(n_r^2 - L)}{(Sn_r^2 - RL)(n_r^2 - P)}, \tag{5}$$

where $n_r = kc/\omega$ is the refractive index, $S = \frac{1}{2}(R + L)$, and

$$\begin{aligned} P &= 1 - \sum_i \frac{\omega_{{\rm p}i}^2}{\omega^2}, \qquad R = 1 - \sum_i \frac{\omega_{{\rm p}i}^2}{\omega^2}\left(\frac{\omega}{\omega + \epsilon_i \Omega_i}\right), \quad \text{and} \\ L &= 1 - \sum_i \frac{\omega_{{\rm p}i}^2}{\omega^2}\left(\frac{\omega}{\omega - \epsilon_i \Omega_i}\right). \end{aligned} \tag{6}$$

Here $\omega_{{\rm p}i}^2 = 4\pi n_i q_i^2/m_i$ and $\Omega_i = |q_i|B/m_i c$ are the plasma and gyro frequencies, $\epsilon_i = q_i/|q_i|$, and n_i, q_i, and m_i are the density, charge, and mass of the background particles. For fully ionised plasmas such as that in the ICM it is sufficient to include terms due to electron, proton and α particles. Figure 1 shows the dispersion surfaces (depicted by the curves) obtained from the above expressions along with the resonant planes in the $(\omega, k_\parallel, k_\perp)$ space. Intersections between the dispersion surfaces and the resonant planes define the resonant wave-particle interactions and the particle kinetic equation coefficients. One can also envision a similar graphic description of the three wave interactions (Eq. 2) using the intersections of the curved dispersion surfaces. However, such calculations have been carried out only in the MHD regime using the simple relations of Eq. 4, which is already a complicated procedure (see e.g. Chandran 2005).

The above dispersion relations are good approximations for low beta plasmas but in the ICM the plasma beta is large:

$$\beta_{\rm p} = 8\pi n {\rm k} T/B^2 = 3.4 \times 10^2 (n/10^{-3}\ {\rm cm}^{-3})(\mu{\rm G}/B)^2 (T/10^8\ {\rm K}). \tag{7}$$

For high beta plasmas the dispersion relation is modified, specially for higher frequencies $\omega \sim k v_{\rm th}$. For example, in the MHD regime, in addition to the Alfvén mode one gets fast and slow modes with the dispersion relation (see e.g. Sturrock 1994)

$$(\omega/k)^2 = \frac{1}{2}\left[(v_{\rm A}^2 + v_{\rm s}^2) \pm \sqrt{v_{\rm A}^4 + v_{\rm s}^4 - 2v_{\rm A}^2 v_{\rm s}^2 \cos 2\theta}\right], \tag{8}$$

and the more general dispersion relation (Eq. 5) is modified in a more complicated way (see e.g. André 1985 or Swanson 1989). The finite temperature imparts an imaginary part $\omega_{\rm i}$ to the wave frequency that gives the damping rate $\Gamma(k)$ as long as $\omega_{\rm i} < \omega_{\rm r}$, the real part of the frequency.[5] For more details see e.g. Barnes and Scargle (1973); Swanson (1989); Pryadko

[5]Note that the 'thermal' effects change $\omega_{\rm r}$ only slightly so that often the real part, the resonant interaction rate and the particle diffusion coefficients can be evaluated using the simpler cold plasma dispersion relation depicted in Fig. 1.

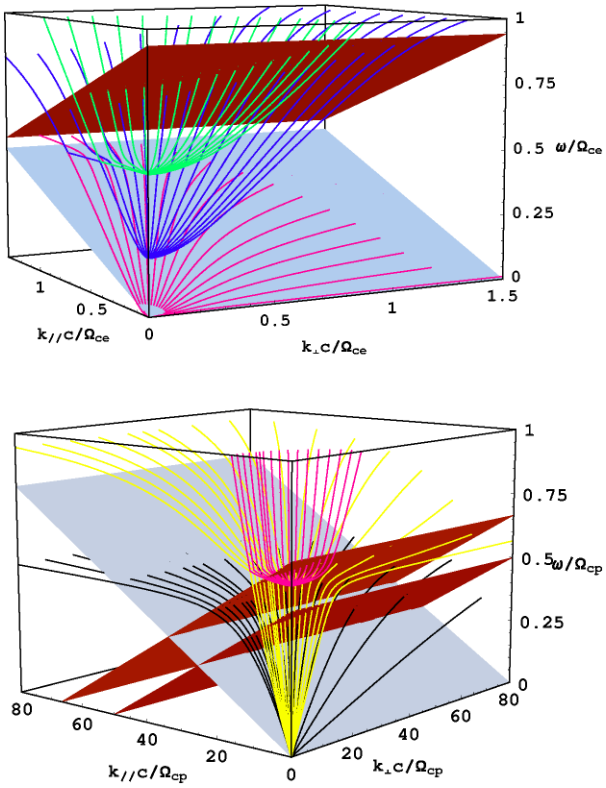


Fig. 1 Dispersion relation (*curves*) surfaces for a *cold* fully ionised H and He (10% by number) plasma and resonance condition (*flat*) surfaces showing the regions around the electron (*top panel*) and proton (*bottom panel*) gyro-frequencies. Only waves with positive $k_\parallel, k_\perp$ (or $0 < \theta < \pi/2$) are shown. The mirror image with respect to the $(\omega, k_\perp)$ plane gives the waves propagating in the opposite direction. From high to low frequencies, we have one of the electromagnetic branches (*green*), upper-hybrid branch (*purple*), lower-hybrid branch, which also includes the whistler waves (*pink*), fast-wave branches (*yellow*), and Alfvén branch (*black*). The effects of a finite temperature modify these curves at frequencies $\omega \sim k v_{\rm th}$, where $v_{\rm th} = \sqrt{2kT/m}$ is the thermal velocity (see e.g. André 1985). The resonance surfaces are for electrons with $v = 0.3$c and $|\mu| = 1.0$ (*top panel*: *upper*, *brown* $n = 1$, *lower*, *light blue* $n = 0$) and ^{4}He (*bottom panel*: *middle*, *brown* $n = 1$) and ^{3}He (*bottom panel*: *upper*, *brown* $n = 1$) ions with $|\mu| = 1.0$ and $v = 0.01$c. The resonance surfaces for the latter two are the same when $n = 0$ (*bottom panel*: *lower*)

and Petrosian (1998, 1999); Cranmer and van Ballegooijen (2003); Brunetti and Lazarian (2007). In general, these rates and the modification of the dispersion relation are known for Maxwellian (sometimes anisotropic) energy distributions of the plasma particles. For non-thermal distributions the damping rates can be evaluated as described Petrosian et al. (2006) using the coupling described in Eq. 11 below.

2.3 Kinetic Equations and Their Coefficients

2.3.1 Wave Equation

Adopting the diffusion approximation (see e.g. Zhou and Matthaeus 1990), one can obtain the evolution of the spatially integrated wave spectrum $W(\mathbf{k}, t)$ from the general equation

$$\frac{\partial W}{\partial t} = \frac{\partial}{\partial k_i}\left[D_{ij}\frac{\partial}{\partial k_j}W\right] - \Gamma(\mathbf{k})W - \frac{W}{T^W_{\rm esc}(\mathbf{k})} + \dot{Q}^W, \tag{9}$$

where $\dot{Q}^W$ is the rate of generation of PWT at $k_{\rm min}$, $T^W_{\rm esc}$ is the escape time, and D_{ij} and Γ describe the cascade and damping of the waves. The calculation of the damping rate is complicated but as described above it is well understood, but there are many uncertainties about the treatment of the cascade process or the form of D_{ij}. This is primarily because of incompleteness of the theoretical models and insufficiency of observational or experimental data. There are some direct observations in the Solar wind (e.g. Leamon et al. 1998) and indirect inferences in the interstellar medium (see e.g. Armstrong et al. 1995). There is some hope (Inogamov and Sunyaev 2003) of future observations in the ICM. Attempts in fitting the Solar wind data have provided some clues about the cascade diffusion coefficients (see Leamon et al. 1999; Jiang et al. 2008).

2.3.2 *Particle Acceleration and Transport*

As described by Petrosian et al. (2008)—Chap. 10, this issue, the general equation for treatment of particles is the Fokker-Planck equation which for ICM conditions can be simplified considerably. As pointed out above we expect a short mean free path and fast scatterings for all particles. When the scattering time $\tau_{\rm scat} = \lambda_{\rm scat}/v \sim \langle 1/D_{\mu\mu}\rangle$ is much less than the dynamic and other timescales, the particles will have an isotropic pitch angle distribution. The pitch-angle averaged and spatially integrated particle distribution is obtained from[6]

$$\frac{\partial N(E,t)}{\partial t} = \frac{\partial}{\partial E}\left[D_{EE}\frac{\partial N(E,t)}{\partial E} - (A - \dot{E}_L)N(E,t)\right] - \frac{N(E,t)}{T^{\rm p}_{\rm esc}} + \dot{Q}^{\rm p}. \tag{10}$$

Here D_{EE}/E^2 is the energy diffusion, due to scattering by PWT as described above and due to Coulomb collisions as discussed by Petrosian et al. (2008)—Chap. 10, this issue, $A(E)/E \sim aD_{EE}/E^2$, with $a(E) = (2-\gamma^{-2})/(1+\gamma^{-1})$ is the rate of direct acceleration due to interactions with PWT and all other agents, e.g., direct first order Fermi acceleration by shocks, $\dot{E}_L/E$ is the energy loss rate of the particles (due to Coulomb collisions and synchrotron and IC losses, see Fig. 4 in Petrosian et al. (2008)—Chap. 10, this issue), and $\dot{Q}^{\rm p}$ and the term with the escape times $T^{\rm p}_{\rm esc}$ describe the source and leakage of particles.[7]

The above two kinetic equations are coupled by the fact that the coefficients of one depend on the spectral distribution of the other; the damping rate of the waves depends on $N(E,t)$ and the diffusion and accelerations rates of particles depend on the wave spectrum $W(\mathbf{k},t)$. Conservation of energy requires that the energy lost by the waves $\dot{\mathcal{W}}_{\rm tot} \equiv \int \Gamma(\mathbf{k})W(\mathbf{k})\mathrm{d}^3k$ must be equal to the energy gained by the particles from the waves; $\dot{\mathcal{E}} = \int [A(E) - A_{\rm sh}]N(E)\mathrm{d}E$. Representing the energy transfer rate between the waves and particles by $\Sigma(\mathbf{k}, E)$ this equality implies that

$$\Gamma(\mathbf{k}) = \int_0^\infty \mathrm{d}E\, N(E)\Sigma(\mathbf{k},E), \qquad A(E) = \int_0^\infty \mathrm{d}^3k W(\mathbf{k})\Sigma(\mathbf{k},E) + A_{\rm sh}, \tag{11}$$

where we have added $A_{\rm sh}$ to represent contributions of other (non-stochastic acceleration) processes affecting the direct acceleration, e.g., shocks.

[6] The derivation of this equation for the stated conditions and some other details can be found in the Appendix.

[7] In what follows we will assume that the waves are confined to the ICM so that $T^W_{\rm esc} \to \infty$ and in some cases we will assume no escape of particles and let $T^{\rm p}_{\rm esc} \to \infty$.

If the damping due to non-thermal particles is important then the wave and particle kinetic equations (9) and (10) are coupled and attempts have been made to obtain solutions of the coupled equations (Miller et al. 1996; Brunetti and Blasi 2005). However, most often the damping rate is dominated by the background thermal particles so that the wave and non-thermal particle kinetic equations decouple. This is a good approximation in the ICM when dealing with relativistic electrons so that for determination of the particle spectra all we need is the boundaries of the inertial range $(k_{\min}, k_{\max})$, the wave spectral index q in this range (most likely $5/3 < q < 3/2$), and the shape of the spectrum above $k_{\max}$ which is somewhat uncertain (see Jiang et al. 2008).

3 Particle Acceleration in Clusters of Galaxies

We now address the problem of particle acceleration in clusters of galaxies. The current information on the ICM does not allow us to treat the problem as outlined above by solving the coupled kinetic equations. In what follows we make reasonable assumptions about the turbulence and the particle diffusion coefficients, and then solve the particle kinetic equation to determine $N(E, t)$. We first consider the apparently simple scenario of acceleration of the background thermal particles. Based on some general arguments, Petrosian (2001, P01 hereafter) showed that this is not a viable mechanism. Here we describe results from a more accurate calculation and show that this indeed is the case. This leads us to consider the transport and acceleration of high energy particles injected into the ICM by other processes.

3.1 Acceleration of Background Particles

The source particles to be accelerated are the ICM hot electrons subject to diffusion in energy space by turbulence and Coulomb collisions, acceleration by turbulence or shocks, and energy losses due to Coulomb collisions.[8] We start with an ICM of $kT = 8$ keV, $n = 10^{-3}$ cm^{-3} and assume a continuous injection of turbulence so that its density remains constant resulting in a time independent diffusion and acceleration rates. The results described below is from a recent paper by Petrosian and East (2007, PE07 hereafter). Following this paper we assume a simple but generic energy dependence of these coefficients. Specifically we assume a simple acceleration rate or timescale

$$\tau_{\rm ac} = E/A(E) = aD(E)/E^2 = \tau_0(1 + E_{\rm c}/E)^p. \tag{12}$$

Figure 2 shows a few examples of these time scales along with the effective Coulomb (plus IC and synchrotron) loss times as described in Fig. 3 of Petrosian et al. (2008)—Chap. 10, this issue.

We then use Eq. 10 to obtain the time evolution of the particle spectra. After each time step we use the resultant spectrum to update the Coulomb coefficients as described by Petrosian et al. (2008)—Chap. 10, this issue. At each step the electron spectrum can be divided into a quasi-thermal and a 'non-thermal' component. A best fit Maxwellian distribution to the quasi-thermal part is obtained, and we determine a temperature and the fraction of the thermal electrons. The remainder is labelled as the non-thermal tail. (For more details see PE07.) The left and middle panels of Fig. 3 show two spectral evolutions for two different

[8] In our numerical results we do include synchrotron, IC and Bremsstrahlung losses. But these have an insignificant effect in the case of nonrelativistic electrons under investigation here.

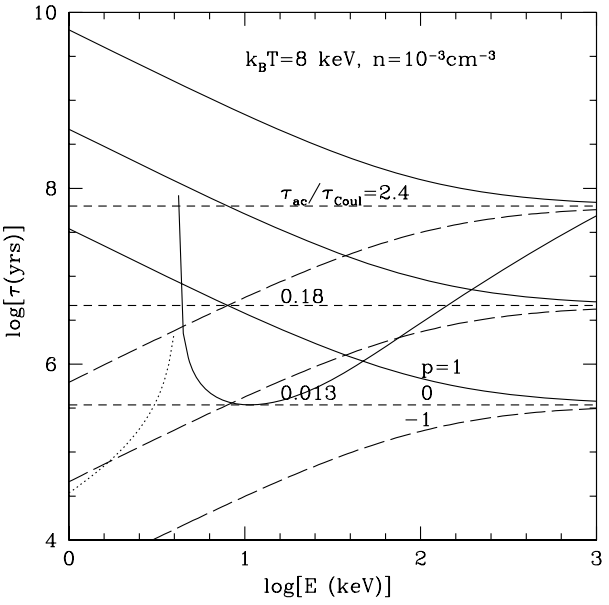


Fig. 2 Acceleration and loss timescales for ICM conditions based on the model described in the text. We use the effective Coulomb loss rate given by Petrosian et al. (2008)—Chap. 10, this issue, and the IC plus synchrotron losses for a CMB temperature of $T_{\rm CMB} = 3$ K and an ICM magnetic field of $B = 1$ μG. We also use the simple acceleration scenario of Eq. 12 for $E_{\rm c} = 0.2 m_{\rm e}c^2$ (~ 100 keV) and for the three specified values of p and times $\tau_0/\tau_{\rm Coul}$ (from Petrosian and East 2007)

values of acceleration time $\tau_0/\tau_{\rm Coul} = 0.013$ and 2.4, respectively, and for $E_{\rm c} = 25$ keV and $p = 1$ ($\tau_{\rm Coul} = \lambda_{\rm Coul}/v = 2.7 \times 10^7$ years). The last spectrum in each case is for time $t = \tau_0$, corresponding to an equal energy input for all cases. The initial and final temperatures, the fraction of particles in the quasi-thermal component $N_{\rm th}$, and the ratio of non-thermal to thermal energies $R_{\rm nonth}$ are shown for each panel. The general feature of these results is that the turbulence causes both acceleration and heating in the sense that the spectra at low energies resemble a thermal distribution but also have a substantial deviation from this quasi-thermal distribution at high energies which can be fitted by a power law over a finite energy range. The distribution is broad and continuous, and as time progresses it becomes broader and shifts to higher energies; the temperature increases and the non-thermal 'tail' becomes more prominent. There is very little evidence for a distinct non-thermal tail for $\tau_0 > \tau_{\rm Coul}$ and most of the turbulent energy goes into heating (top right panel). Note that this also means that for a steady state case where the rate of energy gained from turbulence is equal to radiative energy loss rate (in this case thermal Bremsstrahlung, with time scale $\gg \tau_{\rm Coul}$) there will be an insignificant non-thermal component. There is no distinct non-thermal tail except at unreasonably high acceleration rate (top left panel). Even here there is significant heating (almost doubling of the temperature) within a short time ($\sim 3 \times 10^5$ years). At such rates of acceleration most particles will end up at energies much larger than the initial kT and in a broad non-thermal distribution. We have also calculated spectra for different values of the cutoff energy $E_{\rm c}$ and index p. As expected for larger (smaller) values of $E_{\rm c}$ and smaller (higher) values of p the fraction of non-thermal particles is lower (higher).

The evolution in time of the temperature (in units of its initial value), the fraction of the electrons in the 'non-thermal' component, the energy ratio $R_{\rm nonth}$ as well as an index $\delta = -{\rm d}\ln N(E)/{\rm d}\ln E$ for the non-thermal component are shown in the bottom panel of Fig. 3. All the characteristics described above are more clearly evident in this panel and similar ones for $p = -1$ and $+1$. In all cases the temperature increases by more than a factor of 2. This factor is smaller at higher rates of acceleration. In addition, high acceleration rates

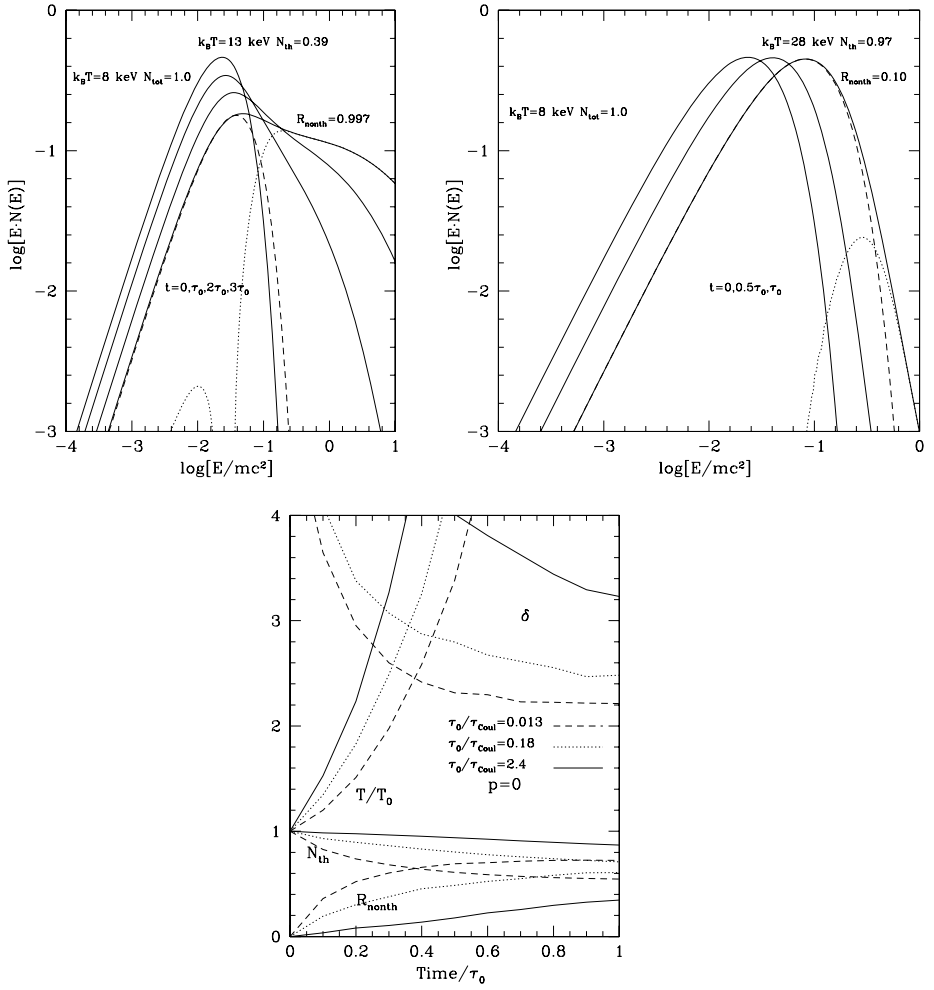


Fig. 3 *Upper left panel*: Evolution with time of electron spectra in the presence of a constant level of turbulence that accelerates electrons according to Eq. 12 with $\tau_0/\tau_{\rm Coul} = 0.013$, $E_{\rm c} = 0.2$ (~ 100 keV) and $p = 1$. For the last spectrum obtained for time $t = \tau_0$, the low end of the spectrum is fitted to a thermal component (*dashed curve*). The residual 'non-thermal' part is shown by the *dotted curve*. We also give the initial and final values of the temperature, the fraction of electrons in the thermal component $N_{\rm th}$, and the ratio of energy of the non-thermal component to the thermal components $R_{\rm nonth}$. *Upper right panel*: Same as above except for $\tau_0/\tau_{\rm Coul} = 2.4$. Note that now there is only heating and not much of acceleration. *Lower panel*: Evolution with time (in units of τ_0) of electron spectral parameters, $T(t)/T_0$, $N_{\rm th}$, $R_{\rm nonth}$ and the power-law index δ for indicated values of $\tau_0/\tau_{\rm Coul}$ and for $p = 0$ and $E_{\rm c} = 100$ keV. Note that for models with the same value of p at $t = \tau_0$ roughly the same amount of energy has been input into the ICM (from Petrosian and East 2007)

produce flatter non-thermal tails (smaller δ) and a larger fraction of non-thermal particles (smaller $N_{\rm th}$) and energy ($R_{\rm nonth}$).

It should be noted that the general aspects of the above behaviour are dictated by the Coulomb collisions and are fairly insensitive to the details of the acceleration mechanism which can affect the spectral evolution somewhat quantitatively but not its qualitative aspects. At low acceleration rates one gets mainly heating and at high acceleration rate

a prominent non-thermal tail is present but there is also substantial heating within one acceleration timescale which for such cases is very short. Clearly in a steady state situation there will be an insignificant non-thermal component. These findings support qualitatively findings by P01 and do not support the presence of distinct non-thermal tails advocated by Blasi (2000) and Dogiel et al. (2007), but agree qualitatively with the more rigorous analysis of Wolfe and Melia (2006). For further results, discussions and comparison with earlier works see PE07.

We therefore conclude that the acceleration of background electrons stochastically or otherwise and non-thermal bremsstrahlung are not a viable mechanism for production of non-thermal hard X-ray excesses observed in some clusters of galaxies.

3.2 Acceleration of Injected Particles

The natural way to overcome the above difficulties is to assume that the radio and the hard X-ray radiation are produced by relativistic electrons injected in the ICM, the first via synchrotron and the second via the inverse Compton scattering of CMB photons. The energy loss rate of relativistic electrons can be approximated by (see P01)

$$\dot{E}_L(E)/E_\mathrm{p} = (1 + (E/E_\mathrm{p})^2)/\tau_\mathrm{loss}, \tag{13}$$

where

$$\tau_\mathrm{loss} = E_\mathrm{p}/(4\pi \mathrm{r}_0^2 m_\mathrm{e}\mathrm{c}^3 n \ln \Lambda) \quad \text{and} \quad E_\mathrm{p} \simeq m_\mathrm{e}\mathrm{c}^2 \left[\frac{9nm_\mathrm{e}\mathrm{c}^2 \ln \Lambda}{8u_\mathrm{ph} + B^2/\pi}\right]^{1/2} \tag{14}$$

are twice the loss time and the energy where the total loss curve reaches its maximum[9] (see Fig. 4). Here $\mathrm{r}_0 = \mathrm{e}^2/(m_\mathrm{e}\mathrm{c}^2) = 2.82 \times 10^{-13}$ cm is the classical electron radius, u_ph and $B^2/8\pi$ are photon (primarily CMB) and magnetic field energy densities. For the ICM $B \sim \mu$G, $n = 10^{-3}$ cm^{-3} and the Coulomb logarithm $\ln \Lambda = 40$ so that $\tau_\mathrm{loss} = 6.3 \times 10^9$ years and $E_\mathrm{p} = 235 m_\mathrm{e}\mathrm{c}^2$.

The electrons are scattered and gain energy if there is some turbulence in the ICM. The turbulence should be such that it resonates with the injected relativistic electrons and not the background thermal nonrelativistic electrons for the reasons described in the previous section. Relativistic electrons will interact mainly with low wavevector waves in the inertial range where $W(k) \propto k^{-q}$ with the index $q \sim 5/3$ or $3/2$ for a Kolmogorov or Kraichnan cascade. There will be little interaction with nonrelativistic background electrons if the turbulence spectrum is cut off above some maximum wave vector k_max whose value depends on viscosity and magnetic field. The coefficients of the transport equation (Eq. 10) can then be approximated by power laws

$$D(E) = \mathcal{D}E^q, \qquad A(E) = a\mathcal{D}E^{q-1}. \tag{15}$$

For a stochastic acceleration model at relativistic energies $a = 2$, but if in addition to scattering by PWT there are other agents of acceleration (e.g. shocks) then the coefficient a will be larger than 2. In this model the escape time is determined by the crossing time $T_\mathrm{cross} \sim R/c$ and the scattering time $\tau_\mathrm{scat} \sim D_{\mu\mu}^{-1}$. We can then write $T_\mathrm{esc} \sim T_\mathrm{cross}(1 + T_\mathrm{cross}/\tau_\mathrm{scat})$. Some examples of these are shown in Fig. 4. However, the escape time is also affected by the

[9] We ignore the Bremsstrahlung loss and the weak dependence on E of Coulomb losses at nonrelativistic energies. We can also ignore the energy diffusion rate due to Coulomb scattering.

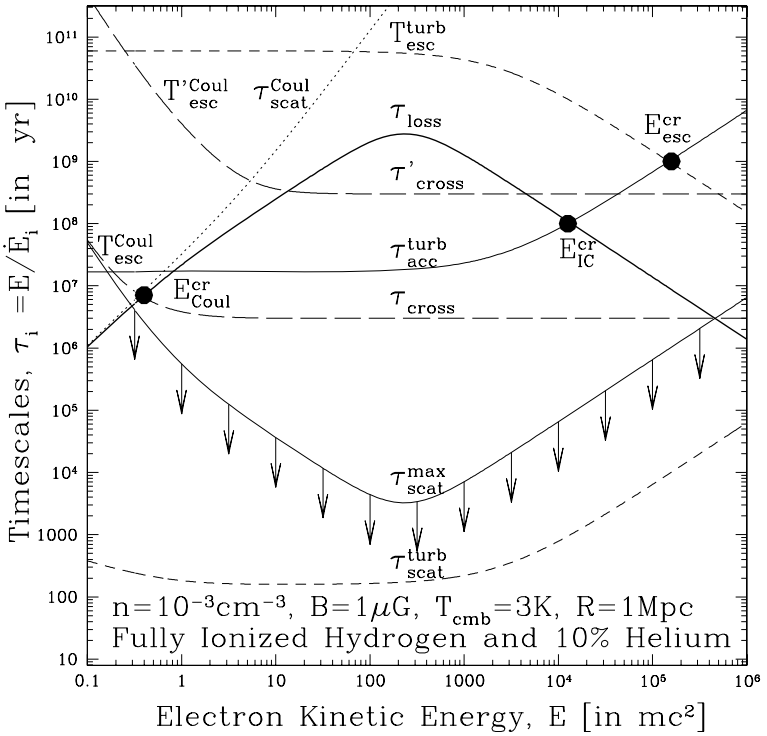


Fig. 4 Comparison of the energy dependence of the total loss time (radiative plus Coulomb; *thick solid line*) with timescales for scattering $\tau_{\rm scat}$ (*lower dashed line*), acceleration $\tau_{\rm ac}$ (*thin solid lines*), crossing time $\tau_{\rm cross} \sim R/c\beta$ (*long dashed lines*) and escape time $T_{\rm esc} \sim \tau^2_{\rm cross}/\tau_{\rm scat}$ (*upper dashed line*) of electrons. At low energies the scattering and escape times are dominated by Coulomb scattering but at high energies turbulence scattering becomes more important. The acceleration is constant at low energies ($A(E) \propto E$) but increases at high energies, corresponding to an acceleration rate $A(E) \rightarrow$ a constant at high E or to a turbulence spectral index $q = 1$. A chaotic magnetic field with scale of 10 kpc will increase the crossing time to the primed curve. The *arrows* show the maximum scattering times for which the escape times are equal or longer than the total loss times. The critical energies where the acceleration time is equal to the escape time, the Coulomb and inverse Compton loss times are shown by *filled circles*

geometry of the magnetic field (e.g. the degree of its entanglement). In what follows we use a power law form $T_{\rm esc} = \mathcal{T}_{\rm esc} E^s$ for the escape time. In addition to these relations we also need the spectrum and rate of injection to obtain the spectrum of radiating electrons. Clearly there are several possibilities. We divide the models into two categories: *steady state* and *time dependent*. In each case we first consider only the effects of losses, which means $\mathcal{D} = 0$ in the above expressions, and then the effects of both acceleration and losses.

3.2.1 Steady State Cases

By steady state we mean variation timescales of order or larger than the Hubble time which is also longer than the maximum loss time $\tau_{\rm loss}/2$. Given a particle injection rate $\dot{Q} = \dot{Q}_0 f(E)$ (with $\int f(E){\rm d}E = 1$) steady state is possible if $\mathcal{T}_{\rm esc} = \dot{Q}_0 / \int N(E) E^{-s} {\rm d}E$.

In the absence of acceleration ($\mathcal{D} = 0$) Eq. 10 can be solved analytically. For the examples of escape times given in Fig. 4 ($T_{\rm esc} > \tau_{\rm loss}$) one gets the simple cooling spectra $N = (\dot{Q}\tau_{\rm loss}/E_{\rm p}) \int_E^\infty f(E){\rm d}E/(1+(E/E_{\rm p})^2)$, which gives a spectral index break at $E_{\rm p}$ from

index $p_0 - 1$ below to $p_0 + 1$ above E_p, for an injected power law $f(E) \propto E^{-p_0}$. For $p_0 = 2$ this will give a high energy power law in rough agreement with the observations but with two caveats. The first is that the spectrum of the injected particles must be cutoff below $E \sim 100 m_e c^2$ to avoid excessive heating and the second is that this scenario cannot produce the broken power law or exponential cutoff we need to explain the radio spectrum of Coma (see Fig. 6 and the discussion in Petrosian et al. (2008)—Chap. 10, this issue). A break is possible only if the escape time is shorter than τ_0 in which case the solution of the kinetic equation for a power law injected spectrum ($p_0 > 1$ and $s > -1$) leads to the broken power law

$$N(E) = Q_0 \begin{cases} \mathcal{T}_{esc}(E/E_p)^{-p_0+s} & \text{if } E \ll E_{cr}, \\ \tau_{loss}(E/E_p)^{-p_0-1}/(s+1) & \text{if } E_{cr} \ll E, \end{cases} \tag{16}$$

where $E_{cr} = E_p((s+1)(\mathcal{T}_{esc}/\tau_{loss})^{-1/(s+1)}$. Thus, for $p_0 \sim 3$ and $s = 0$ and $T_{esc} \simeq 0.02\tau_{loss}$ we obtain a spectrum with a break at $E_{cr} \sim 10^4$, in agreement with the radio data (Rephaeli 1979 model). However, this also means that a large fraction of the $E < E_p$ electrons escape from the ICM, or more accurately from the turbulent confining region, with a flux of $F_{esc}(E) \propto N(E)/T_{esc}(E)$. Such a short escape time means a scattering time which is only ten times shorter than the crossing time and a mean free path of about $\sim 0.1R \sim 100$ kpc. This is in disagreement with the Faraday rotation observations which imply a tangled magnetic field equivalent to a ten times smaller mean free path. The case for a long escape time was first put forth by Jaffe (1977).

Thus it appears that in addition to injection of relativistic electrons we also need a steady presence or injection of PWT to further scatter and accelerate the electrons. The final spectrum of electrons will depend on the acceleration rate and its energy dependence. In general, when the acceleration is dominant one expects a power law spectrum. Spectral breaks appear at critical energies when this rate becomes equal to and smaller than other rates such as the loss or escape rates (see Fig. 4). In the energy range where the losses can be ignored electrons injected at energy E_0 ($f(E) = \delta(E - E_0)$) one expects a power law above (and below, which we are not interested in) this energy. In the realistic case of long T_{esc} (and/or when the direct acceleration rate is larger than the rate of stochastic acceleration (i.e. $a \gg 1$) then spectral index of the electrons will be equal to $-q + 1$ requiring a turbulence spectral index of $q = 4$ which is much larger than expected values of 5/3 or 3/2 (see Park and Petrosian 1995). This spectrum will become steeper (usually cut off exponentially) above the energy where the loss time becomes equal to the acceleration time $\tau_{ac} = E/A(E)$ or at $E_{cr} = (E_p a \mathcal{D} \tau_{loss})^{1/(3-q)}$. Steeper spectra below this energy are possible only for shorter T_{esc}. The left panel of Fig. 5 shows the dependence of the spectra on T_{esc} for $q = 2$ and $s = 0$ (acceleration and escape times independent of E). The spectral index just above E_0 is $p = \sqrt{9/4 + 2\tau_{ac}/T_{esc}} - 1.5$. In the limit when $T_{esc} \to \infty$ the distribution approaches a relativistic Maxwellian distribution $N \propto E^2 e^{-E/E_{cr}}$. For a cut-off energy $E_{cr} \sim 10^4$ this requires an acceleration time of $\sim 10^8$ years and for a spectral index of $p = 3$ below this energy we need $T_{esc} \sim \tau_{ac}/18 \sim 5 \times 10^6$ years which is comparable to the unhindered crossing time. This is too short. As shown in Fig. 4 any scattering mean free path (or magnetic field variation scale) less than the cluster size will automatically give a longer escape time and a flatter than required spectrum. For further detail on all aspects of this case see Park and Petrosian (1995), P01 and Liu et al. (2006).

In summary there are several major difficulties with the steady state model.

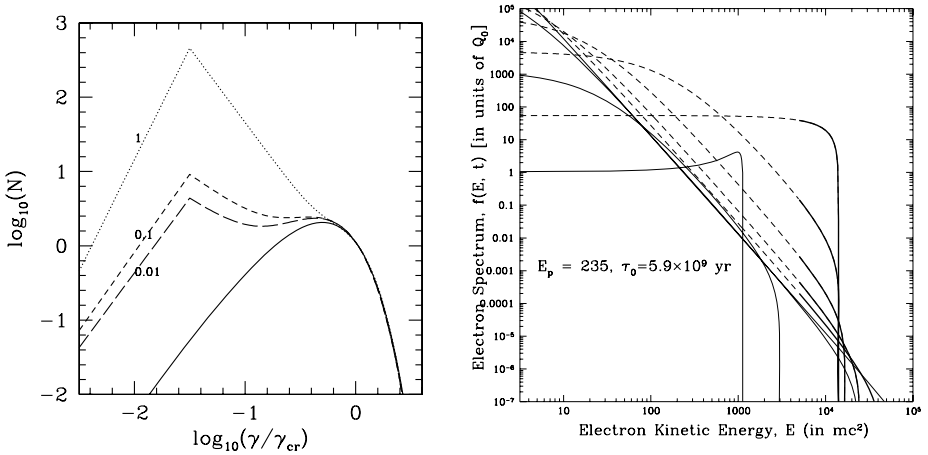


Fig. 5 *Left panel*: The steady state electron spectra injected at energy E_0 and subject to continuous acceleration by turbulence with spectral index $q = 2$ for different values of the ratio $T_{\rm esc}/\tau_{\rm loss}$. For very high values of this ratio we get a relativistic Maxwellian distribution (from Liu et al. 2006). *Right panel*: Evolution with time of a power law injected spectrum (*top line*) subject to Coulomb and inverse Compton (plus synchrotron) losses as given by Eq. 17. *Solid lines* for $b = 2$ ($t/\tau_0 = 10^{-n}$; $n = 3, 2, 1, 0$), and *dashed lines* for $b = 60$ ($t/\tau_0 = 10^{-2.18+n/3}$; $n = 0, 1, 2, 3, 4$). The heavy portions show the energy range of the electrons needed for production of radio and hard X-rays. Note that τ_0 in the label is the same as $\tau_{\rm loss}$ in the text

3.2.2 Time Dependent Models

We are therefore led to consider time dependent scenarios with time variation shorter than the Hubble time. The time dependence may arise from the episodic nature of the injection process (e.g. varying AGN activity) and/or from episodic nature of turbulence generation process (see e.g. Cassano and Brunetti 2005). In this case we need solutions of the time dependent equation (Eq. 10). We start with the generic model of a prompt single-epoch injection of electrons with $Q(E,t) = Q(E)\delta(t - t_0)$. More complex temporal behaviour can be obtained by the convolution of the injection time profile with the solutions described below. The results presented below are from P01. Similar treatments of the following cases can be found in Brunetti et al. (2001) and Brunetti and Lazarian (2007).

It is clear that if there is no re-acceleration, electrons will lose energy first at highest and lowest energies due to inverse Compton and Coulomb losses, respectively. Particles will be peeled away from an initial power law with the low and high energy cut-offs moving gradually toward the peak energy $E_{\rm p}$. A more varied and complex set of spectra can be obtained if we add the effects of diffusion and acceleration. Simple analytic solutions for the time dependent case are possible only for special cases. Most of the complexity arises because of the diffusion term which plays a vital role in shaping the spectrum for a narrow injection spectrum. For some examples see Park and Petrosian (1996). Here we limit our discussion to a broad initial electron spectrum in which case the effects of this term can be ignored until sharp features are developed. Thus, if we set $D(E) = 0$, which is a particularly good approximation when $a \gg 1$, and if for the purpose of demonstration if we again consider the simple case of constant acceleration time ($q = 2$ and $A(E) = a\mathcal{D}E$), then the solution of Eq. 10 gives

$$N(E,t) = \exp\{-t/T_{\rm esc}\}Q_0 \frac{[T_+ - (E/E_{\rm p})\tan(\delta t/\tau_{\rm loss})/\delta]^{p_0-2}}{\cos^2(\delta t/\tau_{\rm loss})[T_-(E/E_{\rm p}) + \tan(\delta t/\tau_{\rm loss})/\delta]^{p_0}}, \qquad (17)$$

where $\delta^2 = 1 - b^2/4$, $b = a\mathcal{D}\tau_0 E_p^2 = \tau_{loss}/\tau_{ac}$ and $T_\pm = 1 \pm b\tan(\delta t/\tau_{loss})/(2\delta)$. Note that $b = 0$ correspond to the case of no acceleration described above. This solution is valid for $b^2 < 4$. For $b^2 > 4$ we are dealing with an imaginary value for δ so that tangents and cosines become hyperbolic functions with $\delta^2 = b^2/4 - 1$. For $\delta = 0$ or $b = 2$ this expression reduces to

$$N(E,t) = \exp\{-t/T_{esc}\} Q_0 \frac{[1 - (E/E_p - 1)t/\tau_{loss}]^{p_0-2}}{[E/E_p - (E/E_p - 1)t/\tau_{loss}]^{p_0}}. \tag{18}$$

The right panel of Fig. 5 shows the evolution of an initial power law spectrum subjected to weak acceleration ($b = 2$, solid lines) and a fairly strong rate of acceleration ($b = 60$, dashed line). As expected with acceleration, one can push the electron spectra to higher levels and extend it to higher energies. At low rates of acceleration the spectrum evolves toward the generic case of a flat low energy part with a fairly steep cutoff above E_p. At higher rates, and for some periods of time comparable to τ_{ac}, the cut off energy E_{cr} will be greater than E_p and there will be a power law portion below it.[10] As evident from this figure there are periods of time when in the relevant energy range (thick solid lines) the spectra resemble what is needed for describing the radio and hard X-ray observations from Coma described in Fig. 6 of Petrosian et al. (2008)—Chap. 10, this issue.

In summary, it appears that a steady state model has difficulties and that the most likely scenario is episodic injection of relativistic particles and/or turbulence and shocks which will re-accelerate the existing or injected relativistic electrons into a spectral shape consistent with observations. However these spectra are short lived, lasting for periods of less than a billion years.

4 Summary and Conclusion

We have given a brief overview of particle acceleration in astrophysical plasmas in general, and acceleration of electrons in the ICM in particular. We have pointed out the crucial role plasma waves and turbulence play in all acceleration mechanisms and outlined the equations that describe the generation, cascade and damping of these waves and the coupling of these processes to the particle kinetics and energising of the plasma and acceleration in both relativistic and nonrelativistic regimes.

We have applied these ideas to the ICM of clusters of galaxies with the aim of production of electron spectra which can explain the claimed hard X-ray emission either as non-thermal bremsstrahlung emission by nonrelativistic electrons or as inverse Compton emission via scattering of CMB photons by a population of relativistic electrons. It is shown that the first possibility which can come about by accelerating background electrons into a non-thermal tail is not a viable mechanism, as was pointed earlier in P01. The primary reason for this difficulty is due to the short Coulomb collision and loss timescales. Quite generally, it can be stated that at low rates of acceleration one obtains a hotter plasma and an insignificant non-thermal tail. Discernible tails can be obtained at higher rates of acceleration but only for very short periods of time. For periods on the order of a billion year such rates will also cause excessive heating and will lead to runaway conditions where most of the electrons are

[10] At even later times than shown here on gets a large pile up at the cut off energy (see P01). This latter feature is of course artificial because we have neglected the diffusion term which will smooth out such features (see Brunetti and Lazarian 2007).

accelerated to relativistic energies, at which they are no longer bound to the cluster, unless there exists a strong scattering agent.

This leads us to the model where hard X-rays are produced by the inverse Compton process and relativistic electrons. Moreover, even if the hard X-ray radiation turns out to be not present, or one finds a way to circumvent the above difficulties, we still require the presence of relativistic electrons to explain the radio emission. These electrons must be injected into the ICM by some other means. They can come from galaxies, specially when they are undergoing an active nuclear (or AGN) phase. Or they may be due to interactions of cosmic ray protons with thermal protons and the resultant pion decays. We have shown that just injection may not be sufficient, because for reasonable injected spectra the transport effects in the ICM modify the spectrum such that the effective radiating spectrum is inconsistent with what is required. Thus, a re-acceleration in the ICM is necessary and turbulence and merger shocks may be the agent of this acceleration. In this case, it also appears that a steady state scenario, like the hadronic mechanism described above, will in general give a flatter than required spectrum unless the electrons escape the ICM unhindered. This requirement is not reasonable because the expected tangled magnetic field will increase this time. But, more importantly, the presence of turbulence necessary for re-acceleration will result in a short mean free path and a much longer escape time.

A more attractive scenario is if the injection of electrons and/or the production of turbulence is episodic. For example for a short lived electron injected phase (from say an AGN) but a longer period of presence of turbulence one can determine the spectral evolution of the electrons subject to acceleration and losses. We have shown that for some periods of time lasting several times the acceleration timescales one can obtain electron spectra consistent with what is required by observations. The same will be true for a hadronic source if there is a short period of production of turbulence. In either case we are dealing with periods on the order of several hundreds of million years to a billion years, which is comparable with timescales expected from merging of Mpc size clusters with velocities of several thousands of $\mathrm{km\,s^{-1}}$ which are theoretically reasonable and agree with observations (see e.g. Bradaĉ et al. 2006).

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". A.M. B. acknowledges the RBRF grant 06-02-16844 and a support from RAS Programs.

appendix

Appendix: Particle Kinetic Equations

In this section we describe some of the mathematical details required for investigation of the acceleration and transport of all charged particles stochastically and by shocks, and the steps and conditions that lead to the specific kinetic equations (Eq. 10) used in this and the previous chapters.

5.1 Stochastic Acceleration by Turbulence

In strong magnetic fields, the gyro-radii of particles are much smaller than the scale of the spatial variation of the field, so that the gyro-phase averaged distribution of the particles depends only on four variables: time, spatial coordinate z along the field lines, the momentum p, and the pitch angle $\cos\mu$. In this case, the evolution of the particle distribution,

$f(t, z, p, \mu)$, can be described by the Fokker-Planck equation as they undergo stochastic acceleration by interaction with plasma turbulence (diffusion coefficients D_{pp}, $D_{\mu\mu}$ and $D_{p\mu}$), direct acceleration (with rate $\dot{p}_G$), and suffer losses (with rate $\dot{p}_L$) due to other interactions with the plasma particles and fields:

$$\frac{\partial f}{\partial t} + v\mu\frac{\partial f}{\partial z} = \frac{1}{p^2}\frac{\partial}{\partial p}p^2\left[D_{pp}\frac{\partial f}{\partial p} + D_{p\mu}\frac{\partial f}{\partial \mu}\right] + \frac{\partial}{\partial \mu}\left[D_{\mu\mu}\frac{\partial f}{\partial \mu} + D_{\mu p}\frac{\partial f}{\partial p}\right] - \frac{1}{p^2}\frac{\partial}{\partial p}[p^2(\dot{p}_L - \dot{p}_G)f] + \dot{J}. \tag{19}$$

Here βc is the velocity of the particles and $\dot{J}(t, z, p, \mu)$ is a source term, which could be the background plasma or some injected spectrum of particles. The kinetic coefficients in the Fokker-Planck equation can be expressed through correlation functions of stochastic electromagnetic fields (see e.g. Melrose 1980; Berezinskii et al. 1990; Schlickeiser 2002). The effect of the mean magnetic field convergence or divergence can be accounted for by adding

$$\frac{c\beta d\ln B}{ds}\frac{\partial}{\partial \mu}\left(\frac{(1-\mu^2)}{2}f\right) \tag{20}$$

to the right hand side.

Pitch-angle isotropy: At high energies and in weakly magnetised plasmas with Alfvén velocity $\beta_A \equiv v_A/c \ll 1$ the ratio of the energy and pitch angle diffusion rates $D_{pp}/p^2 D_{\mu\mu} \approx (\beta_A/\beta)^2 \ll 1$, and one can use the *isotropic approximation* which leads to the *diffusion-convection equation* (see e.g. Dung and Petrosian 1994; Kirk et al. 1988):

$$F(z,t,p) \equiv \frac{1}{2}\int_{-1}^{1} d\mu f(t,z,p,\mu), \qquad \dot{Q}(t,z,p) \equiv \frac{1}{2}\int_{-1}^{1} d\mu \dot{J}(\mu,z,t,p), \tag{21}$$

$$\frac{\partial F}{\partial t} - \frac{\partial}{\partial z}\kappa_1\frac{\partial F}{\partial z} = (pv)\frac{\partial \kappa_2}{\partial z}\frac{\partial F}{\partial p} - \frac{1}{p^2}\frac{\partial}{\partial p}(p^3 v\kappa_2)\frac{\partial F}{\partial z} + \frac{1}{p^2}\frac{\partial}{\partial p}\left(p^4\kappa_3\frac{\partial F}{\partial p} - p^2\dot{p}_L F\right) + \dot{Q}(z,t,p), \tag{22}$$

$$\kappa_1 = \frac{v^2}{8}\int_{-1}^{1} d\mu \frac{(1-\mu^2)^2}{D_{\mu\mu}},$$

$$\kappa_2 = \frac{1}{4}\int_{-1}^{1} d\mu (1-\mu^2)\frac{D_{\mu p}}{pD_{\mu\mu}},$$

$$\kappa_3 = \frac{1}{2}\int_{-1}^{1} d\mu (D_{pp} - D_{\mu p}^2/D_{\mu\mu})p^2.$$

At low energies, as shown by Pryadko and Petrosian (1997), specially for strongly magnetised plasmas ($\alpha \ll 1, \beta_A > 1$), $D_{pp}/p^2 \gg D_{\mu\mu}$, and then stochastic acceleration is more efficient than acceleration by shocks ($D_{pp}/p^2 \gg \dot{p}_G$). In this case the pitch angle dependence may not be ignored.

$$\frac{\partial f^\mu}{\partial t} + v\mu\frac{\partial f^\mu}{\partial z} = \frac{1}{p^2}\frac{\partial}{\partial p}p^2 D_{pp}^{\mu}\frac{\partial f^\mu}{\partial p} - \frac{1}{p^2}\frac{\partial}{\partial p}[p^2\dot{p}_L f^\mu] + \dot{J}^\mu. \tag{23}$$

However, Petrosian and Liu (2004) find that these dependences are in general weak and one can average over the pitch angles.

5.2 Acceleration in Large Scale Turbulence and Shocks

In an astrophysical context it often happens that the energy is released at scales much larger than the mean free path of energetic particles. If the produced large scale MHD turbulence is supersonic and superalfvénic then MHD shocks are present in the system. The particle distribution within such a system is highly intermittent. Statistical description of intermittent systems differs from the description of homogeneous systems. There are strong fluctuations of particle distribution in shock vicinities. A set of kinetic equations for the intermittent system was constructed by Bykov and Toptygin (1993), where the smooth averaged distribution obeys an integro-differential equation (due to strong shocks), and the particle distribution in the vicinity of a shock can be calculated once the averaged function was found.

The pitch-angle averaged distribution function $N(r, p, t)$ of non-thermal particles (with energies below some hundreds of GeV range in the cluster case) averaged over an ensemble of turbulent motions and shocks satisfies the kinetic equation

$$\frac{\partial f}{\partial t} - \frac{\partial}{\partial r_\alpha}\chi_{\alpha\beta}\frac{\partial f}{\partial r_\beta} = G\hat{L}f + \frac{1}{p^2}\frac{\partial}{\partial p}p^4 D\frac{\partial f}{\partial p} + A\hat{L}^2 N + 2B\hat{L}\hat{P}f + \dot{J}(p). \tag{24}$$

The source term $\dot{J}(t, r, p)$ is determined by injection of particles. The integro-differential operators $\hat{L}$ and $\hat{P}$ are given by

$$\hat{L} = \frac{1}{3p^2}\frac{\partial}{\partial p}p^{3-\gamma}\int_0^p \mathrm{d}p' p'^{\gamma}\frac{\partial}{\partial p'}, \qquad \hat{P} = \frac{p}{3}\frac{\partial}{\partial p}. \tag{25}$$

The averaged kinetic coefficients A, B, D, G, and $\chi_{\alpha\beta} = \chi\,\delta_{\alpha\beta}$ are expressed in terms of the spectral functions that describe correlations between large scale turbulent motions and shocks, the particle spectra index γ depends on the shock ensemble properties (see Bykov and Toptygin 1993). The kinetic coefficients satisfy the following renormalisation equations:

$$\chi = \kappa_1(p) + \frac{1}{3}\int \frac{\mathrm{d}^3\mathbf{k}\,\mathrm{d}\omega}{(2\pi)^4}\left[\frac{2T+S}{\mathrm{i}\omega + k^2\chi} - \frac{2k^2\chi S}{(\mathrm{i}\omega + k^2\chi)^2}\right], \tag{26}$$

$$D = \frac{\chi}{9}\int \frac{\mathrm{d}^3\mathbf{k}\,\mathrm{d}\omega}{(2\pi)^4}\frac{k^4 S(k,\omega)}{\omega^2 + k^4\chi^2}, \tag{27}$$

$$A = \chi\int \frac{\mathrm{d}^3\mathbf{k}\,\mathrm{d}\omega}{(2\pi)^4}\frac{k^4\tilde{\phi}(k,\omega)}{\omega^2 + k^4\chi^2}, \tag{28}$$

$$B = \chi\int \frac{\mathrm{d}^3\mathbf{k}\,\mathrm{d}\omega}{(2\pi)^4}\frac{k^4\tilde{\mu}(k,\omega)}{\omega^2 + k^4\chi^2}. \tag{29}$$

Here $G = (1/\tau_{\mathrm{sh}} + B)$. $T(k,\omega)$ and $S(k,\omega)$ are the transverse and longitudinal parts of the Fourier components of the turbulent velocity correlation tensor. Correlations between velocity jumps on shock fronts are described by $\tilde{\phi}(k,\omega)$, while $\tilde{\mu}(k,\omega)$ represents shock-rarefaction correlations. The introduction of these spectral functions is dictated by the intermittent character of a system with shocks.

The test particle calculations showed that the low energy branch of the particle distribution would contain a substantial fraction of the free energy of the system after a few acceleration times. Thus, to calculate the efficiency of the shock turbulence power conversion

to the non-thermal particle component, as well as the particle spectra, we have to account for the backreaction of the accelerated particles on the shock turbulence. To do that, Bykov (2001) supplied the kinetic equations Eqs. 24–29 with the energy conservation equation for the total system including the shock turbulence and the non-thermal particles, resulting in temporal evolution of particle spectra.

References

A.I. Akhiezer, I.A. Akhiezer, R.V. Polovin, A.G. Sitenko, K.N. Stepanov, *Plasma Electrodynamics* (Pergamon, Elmsford, 1975)
E. Amato, P. Blasi, Mon. Not. R. Astron. Soc. **371**, 1251 (2006)
M. André, J. Plasma Phys. **33**, 1 (1985)
J.W. Armstrong, B.J. Rickett, S.R. Spangler, Astrophys. J. **443**, 209 (1995)
A. Barnes, J.D. Scargle, Astrophys. J. **184**, 251 (1973)
A.R. Bell, S.G. Lucek, Mon. Not. R. Astron. Soc. **321**, 433 (2001)
V.S. Berezinskii, S.V. Bulanov, V.A. Dogiel, V.S. Ptuskin, *Astrophysics of Cosmic Rays* (North-Holland, Amsterdam, 1990)
R.D. Blandford, D. Eichler, Phys. Rep. **154**, 1 (1987)
R.D. Blandford, J.P. Ostriker, Astrophys. J. **221**, L29 (1978)
P. Blasi, Astrophys. J. **532**, L9 (2000)
J.P. Boris, J.M. Dawson, J.H. Orens, K.V. Roberts, Phys. Rev. Lett. **25**, 706 (1970)
M. Bradaĉ, D. Clowe, A.H. Gonzalez et al., Astrophys. J. **652**, 937 (2006)
G. Brunetti, P. Blasi, Mon. Not. R. Astron. Soc. **363**, 1173 (2005)
G. Brunetti, A. Lazarian, Mon. Not. R. Astron. Soc. **378**, 245 (2007)
G. Brunetti, G. Setti, L. Feretti, G. Giovannini, Mon. Not. R. Astron. Soc. **320**, 365 (2001)
A.M. Bykov, Space Sci. Rev. **99**, 317 (2001)
A.M. Bykov, I.N. Toptygin, Phys. Uspekhi **36**, 1020 (1993)
A.M. Bykov, H. Bloemen, Yu.A. Uvarov, Astron. Astrophys. **362**, 886 (2000)
A.M. Bykov, K. Dolag, F. Durret, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9312-9
P.A. Cassak, J.F. Drake, M.A. Shay, Astrophys. J. **644**, L145 (2006)
R. Cassano, G. Brunetti, Mon. Not. R. Astron. Soc. **357**, 1313 (2005)
B. Chandran, Phys. Rev. Lett. **95**, 265004 (2005)
J. Cho, A. Lazarian, Phys. Rev. Lett. **88**, 245001 (2002)
J. Cho, A. Lazarian, Astrophys. J. **638**, 811 (2006)
S.R. Cranmer, A.A. van Ballegooijen, Astrophys. J. **594**, 573 (2003)
V.A. Dogiel, C.M. Ko, P.H. Kuo et al., Astron. Astrophys. **461**, 433 (2007)
K. Dolag, F. Vazza, G. Brunetti, G. Tormen, Mon. Not. R. Astron. Soc. **364**, 753 (2005)
J.F. Drake, Paper presented at Krakow Conference on Relativistic Jets, http://www.oa.uj.edu.pl/2006jets/talks/Drake/Drake.pdf (2006)
L.Oc. Drury, Rep. Prog. Phys. **46**, 973 (1983)
R. Dung, V. Petrosian, Astrophys. J. **421**, 550 (1994)
F. Durret, J.S. Kaastra, J. Nevalainen, T. Ohashi, N. Werner, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9313-8
D.C. Ellison, E. Moebius, G. Paschmann, Astrophys. J. **352**, 376 (1990)
D.C. Ellison, A. Decourchelle, J. Ballet, Astron. Astrophys. **429**, 569 (2005)
C. Ferrari, F. Govoni, S. Schindler, A.M. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9311-x
P. Goldreich, S. Sridhar, Astrophys. J. **438**, 763 (1995)
P. Goldreich, S. Sridhar, Astrophys. J. **485**, 680 (1997)
G.D. Holman, Astrophys. J. **293**, 584 (1985)
N.A. Inogamov, R.A. Sunyaev, Astron. Lett. **29**, 791 (2003)
W.J. Jaffe, Astrophys. J. **212**, 1 (1977)
Y.W. Jiang et al., Astrophys. J. (2008, submitted). arXiv:0802.0910
F.C. Jones, Astrophys. J. Suppl. Ser. **90**, 561 (1994)
F.C. Jones, D.C. Ellison, Space Sci. Rev. **58**, 259 (1991)
C.F. Kennel, F.V. Coroniti, F.L. Scarf et al., J. Geophys. Res. **91**, 11917 (1986)
J.G. Kirk, P. Schneider, R. Schlickeiser, Astrophys. J. **328**, 269 (1988)
R.J. Leamon, C.W. Smith, N.F. Ness, W.H. Matthaeus, H.K. Wong, J. Geophys. Res. **103**, 4775 (1998)

R.J. Leamon, C.W. Smith, N.F. Ness, H.K. Wong, J. Geophys. Res. **104**, 22331 (1999)
Y. Lithwick, P. Goldreich, Astrophys. J. **582**, 1220 (2003)
S. Liu, F. Melia, V. Petrosian, Astrophys. J. **636**, 798 (2006)
M.A. Malkov, L.Oc. Drury, Rep. Prog. Phys. **64**, 429 (2001)
D.B. Melrose, *Plasma Astrophysics* (Gordon and Breach, New York, 1980)
J.A. Miller, T.N. LaRosa, R.L. Moore, Astrophys. J. **461**, 445 (1996)
B.T. Park, V. Petrosian, Astrophys. J. **446**, 699 (1995)
B.T. Park, V. Petrosian, Astrophys. J. Suppl. Ser. **103**, 255 (1996)
V. Petrosian, Astrophys. J. **557**, 560 (2001) (P01)
V. Petrosian, W. East, Astrophys. J. (2007, in press) (PE07). arXiv:0802.0900
V. Petrosian, S. Liu, Astrophys. J. **610**, 550 (2004)
V. Petrosian, H. Yan, A. Lazarian, Astrophys. J. **644**, 603 (2006)
V. Petrosian, A.M. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9327-2
J. Pryadko, V. Petrosian, Astrophys. J. **482**, 774 (1997)
J. Pryadko, V. Petrosian, Astrophys. J. **495**, 377 (1998)
J. Pryadko, V. Petrosian, Astrophys. J. **515**, 873 (1999)
Y. Rephaeli, Astrophys. J. **227**, 364 (1979)
Y. Rephaeli, J. Nevalainen, T. Ohashi, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9314-7
P.M. Ricker, C.L. Sarazin, Astrophys. J. **561**, 621 (2001)
K. Roettiger, J.O. Burns, C. Loken, Astrophys. J. **473**, 561 (1996)
A.A. Schekochihin, S.C. Cowley, R.M. Kulsrud, G.W. Hammett, P. Sharma, Astrophys. J. **626**, 139 (2005)
R. Schlickeiser, *Cosmic Ray Astrophysics* (Springer, Berlin, 2002)
R. Schlickeiser, A. Campeanu, L. Lerche, Astron. Astrophys. **276**, 614 (1993)
R.A. Sunyaev, M.L. Norman, G.L. Bryan, Astron. Lett. **29**, 783 (2003)
D.G. Swanson, *Plasma Waves* (Academic Press, New York, 1989)
P.A. Sturrock, *Plasma Physics* (Cambridge Univ. Press, Cambridge, 1994)
A. Vladimirov, D.C. Ellison, A. Bykov, Astrophys. J. **652**, 1246 (2006)
C. Vogt, T.A. Enßlin, Astron. Astrophys. **434**, 67 (2005)
H.J. Völk, E.G. Berezhko, L.T. Ksenofontov, Astron. Astrophys. **433**, 229 (2005)
B. Wolfe, F. Melia, Astrophys. J. **638**, 125 (2006)
S. Zenitani, M. Hoshino, Astrophys. J. **618**, L111 (2005)
Y. Zhou, W.H. Matthaeus, J. Geophys. Res. **95**, 14881 (1990)

Chapter 12
Simulation Techniques for Cosmological Simulations

K. Dolag · S. Borgani · S. Schindler · A. Diaferio · A.M. Bykov

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9316-5

Abstract Modern cosmological observations allow us to study in great detail the evolution and history of the large scale structure hierarchy. The fundamental problem of accurate constraints on the cosmological parameters, within a given cosmological model, requires precise modelling of the observed structure. In this paper we briefly review the current most effective techniques of large scale structure simulations, emphasising both their advantages and shortcomings. Starting with basics of the direct N-body simulations appropriate to modelling cold dark matter evolution, we then discuss the direct-sum technique *GRAPE*, particle-mesh (*PM*) and hybrid methods, combining the *PM* and the tree algorithms. Simulations of baryonic matter in the Universe often use hydrodynamic codes based on both particle methods that discretise mass, and grid-based methods. We briefly describe Eulerian grid methods, and also some variants of Lagrangian smoothed particle hydrodynamics (*SPH*) methods.

K. Dolag (✉)
Max-Planck-Institut für Astrophysik, P.O. Box 1317, 85741 Garching, Germany
e-mail: kdolag@mpa-garching.mpg.de

S. Borgani
Department of Astronomy, University of Trieste, via Tiepolo 11, 34143 Trieste, Italy
e-mail: borgani@oats.inaf.it

S. Schindler
Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria
e-mail: sabine.schindler@uibk.ac.at

A. Diaferio
Dipartimento di Fisica Generale "Amedeo Avogadro", Università degli Studi di Torino, Turin, Italy
e-mail: diaferio@ph.unito.it

A. Diaferio
Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Torino, via P. Giuria 1, 10125 Turin, Italy

A.M. Bykov
A.F. Ioffe Institute of Physics and Technology, 194021 St. Petersburg, Russia
e-mail: byk@astro.ioffe.ru

Keywords Cosmology: theory · Large-scale structure of universe · Hydrodynamics · Method: numerical, N-body simulations

1 Introduction

In the hierarchical picture of structure formation, small objects collapse first and then merge to form larger and larger structures in a complex manner. This formation process reflects on the intricate structure of galaxy clusters, whose properties depend on how the thousands of smaller objects that the cluster accretes are destroyed or survive within the cluster gravitational potential. These merging events are the source of shocks, turbulence and acceleration of relativistic particles in the intracluster medium, which, in turn, lead to a redistribution or amplification of magnetic fields, and to the acceleration of cosmic rays. In order to model these processes realistically, we need to resort to numerical simulations which are capable of resolving and following correctly the highly non-linear dynamics. In this paper, we briefly describe the methods which are commonly used to simulate galaxy clusters within a cosmological context.

Usually, choosing the simulation setup is a compromise between the size of the region that one has to simulate to fairly represent the object(s) of interest, and the resolution needed to resolve the objects at the required level of detail. Typical sizes of the simulated volume are a megaparsec scale for an individual galaxy, tens to hundreds of megaparsecs for a galaxy population, and several hundreds of megaparsecs for a galaxy cluster population. The mass resolution varies from $\approx 10^5$ $\mathrm{M}_\odot$ up to $\approx 10^{10}$ $\mathrm{M}_\odot$, depending on the object studied, while, nowadays, one can typically reach the resolution of a few hundred parsec for individual galaxies and above the kiloparsec scale for cosmological boxes.

2 N-Body (Pure Gravity)

Over most of the cosmic time of interest for structure formation, the Universe is dominated by dark matter. The most favourable model turned out to be the so-called cold dark matter (CDM) model. The CDM can be described as a collisionless, non-relativistic fluid of particles of mass m, position $\mathbf{x}$ and momentum $\mathbf{p}$. In an expanding background Universe (usually described by a Friedmann-Lemaître model), with $a = (1 + z)^{-1}$ being the Universe scale factor, $\mathbf{x}$ is the comoving position and the phase-space distribution function $f(\mathbf{x}, \mathbf{p}, t)$ of the dark-matter fluid can be described by the collisionless Boltzmann (or Vlasov) equation

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{ma^2}\nabla f - m\nabla\Phi\frac{\partial f}{\partial \mathbf{p}} = 0 \tag{1}$$

coupled with the Poisson equation

$$\nabla^2\Phi(\mathbf{x}, t) = 4\pi \mathrm{G}a^2\left[\rho(\mathbf{x}, t) - \bar{\rho}(t)\right], \tag{2}$$

where Φ is the gravitational potential and $\bar{\rho}(t)$ is the background density. The proper mass density

$$\rho(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{p}, t)\mathrm{d}^3 p \tag{3}$$

can be inferred by integrating the distribution function over the momenta $\mathbf{p} = ma^2\dot{\mathbf{x}}$.

This set of equations represents a high-dimensional problem. It is therefore usually solved by sampling the phase-space density by a finite number N of tracer particles. The solution can be found through the equation of motion of the particles (in comoving coordinates),

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -m\nabla\Phi \tag{4}$$

and

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{\mathbf{p}}{ma^2}. \tag{5}$$

Introducing the proper peculiar velocity $\mathbf{v} = a\dot{\mathbf{x}}$ these equations can be written as

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + \mathbf{v}\frac{\dot{a}}{a} = -\frac{\nabla\Phi}{a}. \tag{6}$$

The time derivative of the expansion parameter, $\dot{a}$, can be obtained from the Friedmann equation

$$\dot{a} = H_0\sqrt{1 + \Omega_0(a^{-1} - 1) + \Omega_\Lambda(a^2 - 1)}, \tag{7}$$

where we have assumed the dark energy to be equivalent to a cosmological constant. For a more detailed description of the underlying cosmology and related issues, see for example Peebles (1980) or others.

There are different approaches: to solve directly the motion of the tracer particles, or to solve the Poisson equation. Some of the most common methods will be described briefly in the following sections.

2.1 Direct Sum (GRAPE, GPU)

The most direct way to solve the N-body problem is to sum directly the contributions of all the individual particles to the gravitational potential

$$\Phi(\mathbf{r}) = -\mathrm{G}\sum_j \frac{m_j}{(|\mathbf{r} - \mathbf{r}_j|^2 + \epsilon^2)^{\frac{1}{2}}}. \tag{8}$$

In principle, this sum would represent the exact (Newtonian) potential which generates the particles' acceleration. As mentioned before, the particles do not represent individual dark matter particles, but should be considered as Monte Carlo realisations of the mass distribution, and therefore only collective, statistical properties can be considered. In such simulations, close encounters between individual particles are irrelevant to the physical problem under consideration, and the gravitational force between two particles is smoothed by introducing the gravitational softening ϵ. This softening reduces the spurious two-body relaxation which occurs when the number of particles in the simulation is not large enough to represent correctly a collisionless fluid. This situation however is unavoidable, because the number of dark matter particles in real systems is orders of magnitude larger than the number that can be handled in a numerical simulation. Typically, ϵ is chosen to be 1/20–1/50 of the mean inter-particle separation within the simulation. In general, this direct-sum approach is considered to be the most accurate technique, and is used for problems where superior precision is needed. However this method has the disadvantage of being already quite CPU intensive for even a moderate number of particles, because the computing time is $\propto N^2$, where N is the total number of particles.

Rather than searching for other software solutions, an alternative approach to solve the N^2-bottleneck of the direct-sum technique is the *GRAPE* (GRAvity PipE) special-purpose hardware (see e.g. Ito et al. 1993 and related articles). This hardware is based on custom chips that compute the gravitational force with a hardwired Plummer force law (8). This hardware device thus solves the gravitational N-body problem with a direct summation approach at a computational speed which is considerably higher than that of traditional processors.

For the force computation, the particle coordinates are first loaded onto the *GRAPE* board, then the forces for several positions (depending on the number of individual *GRAPE* chips installed in the system) are computed in parallel. In practise, there are some technical complications when using the *GRAPE* system. One is that the hardware works internally with special fixed-point formats or with limited floating point precision (depending on the version of the *GRAPE* chips used) for positions, accelerations and masses. This results in a reduced dynamic range compared to the standard IEEE floating point arithmetic. Furthermore, the communication time between the host computer and the *GRAPE* system can be an issue in certain circumstances. However, newer versions of the *GRAPE* chips circumvent this problem, and can also be combined with the tree algorithms (which are described in detail in the next section), see Fukushige et al. (1991), Makino (1991), Athanassoula et al. (1998), Kawai et al. (2000).

By contrast, the graphic processing unit (*GPU*) on modern graphic cards now provides an alternative tool for high-performance computing. The original purpose of the *GPU* is to serve as a graphics accelerator for speeding up the image processing, thereby allowing one to perform simple instructions on multiple data. It has therefore become an active area of research to use the *GPU*s of the individual members of computer clusters. Although very specialised, many of those computational algorithms are also needed in computational astrophysics, and therefore the *GPU* can provide significantly more computing power than the host system; thereby providing a high performance with typically large memory size and at relatively low cost, which represents a valid alternative to special purpose hardware like *GRAPE*. For recent applications to astrophysical problems see Schive et al. (2007) and references therein.

2.2 Tree

The primary method of solving the N-body problem is a hierarchical multipole expansion, commonly called a tree algorithm. This method groups distant particles into larger cells, allowing their gravity to be accounted for by means of a single multipole force. Instead of requiring $N-1$ partial force evaluations per particle, as needed in a direct-summation approach, the gravitational force on a single particle can be computed with substantially fewer operations, because distant groups are treated as "macro" particles in the sum. In this manner the sum usually reduces to $N\log(N)$ operations. Note however that this scaling is only true for homogeneous particle distributions, whereas the scaling for strongly inhomogeneous distributions, as present in evolved cosmological structures, can be less efficient.

In practise, the hierarchical grouping that forms the basis of the multipole expansion is most commonly obtained by a recursive subdivision of space. In the approach of Barnes and Hut (1986), a cubical root node is used to encompass the full mass distribution; the cube is repeatedly subdivided into eight daughter nodes of half the side-length each, until one ends up with 'leaf' nodes containing single particles (see Fig. 1). Forces are then obtained by "walking" the tree. In other words, starting at the root node, a decision is made as to whether or not the multipole expansion of the node provides an accurate enough partial

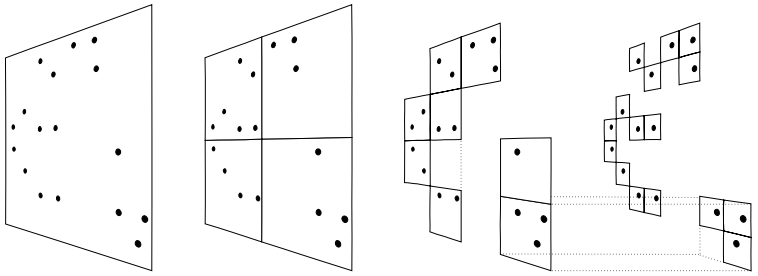


Fig. 1 Schematic illustration of the Barnes and Hut (1986) oct-tree in two dimensions. The particles are first enclosed in a square (root node). This square is then iteratively subdivided into four squares of half the size, until exactly one particle is left in each final square (leaves of the tree). In the resulting tree structure, each square can be the progenitor of up to four siblings. Taken from Springel et al. (2001b)

force. If the answer is 'yes', the multipole force is used and the walk along this branch of the tree can be terminated; if the answer is 'no', the node is "opened", i.e. its daughter nodes are considered in turn. Clearly, the multipole expansion is in general appropriate for nodes that are sufficiently small and distant. Most commonly one uses a fixed angle (typically ≈ 0.5 rad) as opening criteria.

It should be noted that the final result of the tree algorithm will in general only represent an approximation to the true force. However, the error can be controlled conveniently by modifying the opening criterion for tree nodes, because a higher accuracy is obtained by walking the tree to lower levels. Provided that sufficient computational resources are invested, the tree force can then be made arbitrarily close to the well-specified correct force. Nevertheless evaluating the gravitational force via a tree leads to an inherent asymmetry in the interaction between two particles. It is worth mentioning that there are extensions to the standard tree, the so-called *fast multipole methods*, which avoid these asymmetries, and therefore have better conservation of momentum. For an N-body application of such a technique see Dehnen (2000) and references therein. However, these methods compute the forces for all the particles at every time step and can not take advantage of using individual time steps for different particles.

2.3 Particle-Mesh Methods

The Particle-Mesh (*PM*) method treats the force as a field quantity by computing it on a mesh. Differential operators, such as the Laplacian, are replaced by finite difference approximations. Potentials and forces at particle positions are obtained by interpolation on the array of mesh-defined values. Typically, such an algorithm is performed in three steps. First, the density on the mesh points is computed by assigning densities to the mesh from the particle positions. Second, the density field is transformed to Fourier space, where the Poisson equation is solved, and the potential is obtained using Green's method. Alternatively, the potential can be determined by solving Poisson's equation iteratively with relaxation methods. In a third step the forces for the individual particles are obtained by interpolating the derivatives of the potentials to the particle positions. Typically, the amount of mesh cells N used corresponds to the number of particles in the simulation, so that when structures form, one can have large numbers of particles within individual mesh cells, which immediately illustrates the shortcoming of this method; namely its limited resolution. On the other hand, the calculation of the Fourier transform via a Fast Fourier Transform (FFT) is extremely fast, as

it only needs of order $N \log N$ operations, which is the advantage of this method. Note that here N denotes the number of mesh cells. In this approach the computational costs do not depend on the details of the particle distribution. Also this method can not take advantage of individual time steps, as the forces are always calculated for all particles at every time step.

There are many schemes to assign the mass density to the mesh. The simplest method is the "Nearest-Grid-Point" (NGP). Here, each particle is assigned to the closest mesh point, and the density at each mesh point is the total mass assigned to the point divided by the cell volume. However, this method is rarely used. One of its drawbacks is that it gives forces that are discontinuous. The "Cloud-in-a-Cell" (CIC) scheme is a better approximation to the force: it distributes every particle over the nearest 8 grid cells, and then weighs them by the overlapping volume, which is obtained by assuming the particle to have a cubic shape of the same volume as the mesh cells. The CIC method gives continuous forces, but discontinuous first derivatives of the forces. A more accurate scheme is the "Triangular-Shaped-Cloud" (TSC) method. This scheme has an assignment interpolation function that is piecewise quadratic. In three dimensions it employs 27 mesh points (see Hockney and Eastwood 1988).

In general, one can define the assignment of the density ρ_m on a grid $\mathbf{x}_m$ with spacing δ from the distribution of particles with masses m_i and positions $\mathbf{x}_i$, by smoothing the particles over n times the grid spacing ($h = n\delta$). Therefore, having defined a weighting function

$$W(\mathbf{x}_m - \mathbf{x}_i) = \int \hat{W}\left(\frac{\mathbf{x} - \mathbf{x}_m}{h}\right) S(\mathbf{x} - \mathbf{x}_i, h)\mathrm{d}\mathbf{x}, \tag{9}$$

where $\hat{W}(\mathbf{x})$ is 1 for $|\mathbf{x}| < 0.5$ and 0 otherwise, the density ρ_m on the grid can be written as

$$\rho_m = \frac{1}{h^3} \sum_i m_i W(\mathbf{x}_i - \mathbf{x}_m). \tag{10}$$

The shape function $S(\mathbf{x}, h)$ then defines the different schemes. The aforementioned NGP, CIC and TSC schemes are equivalent to the choice of 1, 2 or 3 for n and the Dirac δ function $\delta(\mathbf{x})$, $\hat{W}(\mathbf{x}/h)$ and $1 - |\mathbf{x}/h|$ for the shape function $S(\mathbf{x}, h)$, respectively.

In real space, the gravitational potential Φ can be written as the convolution of the mass density with a suitable Green's function $g(\mathbf{x})$:

$$\Phi(\mathbf{x}) = \int g(\mathbf{x} - \mathbf{x}')\rho(\mathbf{x}')\mathrm{d}\mathbf{x}'. \tag{11}$$

For vacuum boundary conditions, for example, the gravitational potential is

$$\Phi(\mathbf{x}) = -\mathrm{G} \int \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}\mathrm{d}\mathbf{x}', \tag{12}$$

with G being the gravitational constant. Therefore the Green's function, $g(\mathbf{x}) = -\mathrm{G}/|\mathbf{x}|$, represents the solution of the Poisson equation $\nabla^2\Phi(\mathbf{x}) = 4\pi\mathrm{G}\rho(\mathbf{x})$, recalling that $\nabla_x^2(|\mathbf{x} - \mathbf{x}'|)^{-1} = 4\pi\delta(\mathbf{x} - \mathbf{x}')$. By applying the divergence theorem to the integral form of the above equation, it is then easy to see that, in spherical coordinates,

$$\int_V \nabla^2\left(\frac{1}{r}\right)\mathrm{d}V = \int_S \nabla\left(\frac{1}{r}\right)\mathrm{d}S = \int_0^{2\pi}\int_0^{\pi} \frac{\partial}{\partial r}\left(\frac{1}{r}\right) r^2 \sin(\theta)\mathrm{d}\theta\mathrm{d}\phi = -4\pi. \tag{13}$$

Periodic boundary conditions are usually used to simulate an "infinite universe", however zero padding can be applied to deal with vacuum boundary conditions.

In the *PM* method, the solution to the Poisson equation is easily found in Fourier space, where (11) becomes a simple multiplication

$$\hat{\Phi}(\mathbf{k}) = \hat{g}(\mathbf{k})\,\hat{\rho}(\mathbf{k}). \tag{14}$$

Note that $\hat{g}(\mathbf{k})$ has only to be computed once, at the beginning of the simulation.

After the calculation of the potential via Fast Fourier Transform (FFT) methods, the force field $\mathbf{f}(\mathbf{x})$ at the position of the mesh points can be obtained by differentiating the potential, $\mathbf{f}(\mathbf{x}) = \nabla\Phi(\mathbf{x})$. This can be done by a finite-difference representation of the gradient. In a second order scheme, the derivative with respect to the x coordinate at the mesh positions $m = (i, j, k)$ can be written as

$$f^{(x)}_{i,j,k} = -\frac{\Phi_{i+1,j,k} - \Phi_{i-1,j,k}}{2h}. \tag{15}$$

A fourth order scheme for the derivative would be written as

$$f^{(x)}_{i,j,k} = -\frac{4}{3}\frac{\Phi_{i+1,j,k} - \Phi_{i-1,j,k}}{2h} + \frac{1}{3}\frac{\Phi_{i+2,j,k} - \Phi_{i-2,j,k}}{4h}. \tag{16}$$

Finally, the forces have to be interpolated back to the particle positions as

$$\mathbf{f}(\mathbf{x}_i) = \sum_m W(\mathbf{x}_i - \mathbf{x}_m)\mathbf{f}_m, \tag{17}$$

where it is recommended to use the same weighting scheme as for the density assignment; this ensures pairwise force symmetry between particles and momentum conservation.

The advantage of such *PM* methods is the speed, because the number of operations scales with $N + N_g \log(N_g)$, where N is the number of particles and N_g the number of mesh points. However, the disadvantage is that the dynamical range is limited by N_g, which is usually limited by the available memory. Therefore, particularly for cosmological simulations, adaptive methods are needed to increase the dynamical range and follow the formation of individual objects.

In the Adaptive Mesh Refinement (*AMR*) techniques, the Poisson equation on the refinement meshes can be treated as a Dirichlet boundary problem for which the boundary values are obtained by interpolating the gravitational potential from the parent grid. In such algorithms, the boundaries of the refinement meshes can have an arbitrary shape; this feature narrows the range of solvers that one can use for partial differential equation (PDEs). The Poisson equation on these meshes can be solved using the *relaxation* method (Hockney and Eastwood 1988; Press et al. 1992), which is relatively fast and efficient in dealing with complicated boundaries. In this method the Poisson equation

$$\nabla^2\Phi = \rho \tag{18}$$

is rewritten in the form of a diffusion equation,

$$\frac{\partial\Phi}{\partial\tau} = \nabla^2\Phi - \rho. \tag{19}$$

The point of the method is that an initial solution guess Φ *relaxes* to an equilibrium solution (i.e., solution of the Poisson equation) as $\tau \to \infty$. The finite-difference form of (2) is:

$$\Phi^{n+1}_{i,j,k} = \Phi^n_{i,j,k} + \frac{\Delta\tau}{\Delta^2}\left(\sum_{nb=1}^{6}\Phi^n_{nb} - 6\Phi^n_{i,j,k}\right) - \rho_{i,j,k}\Delta\tau, \tag{20}$$

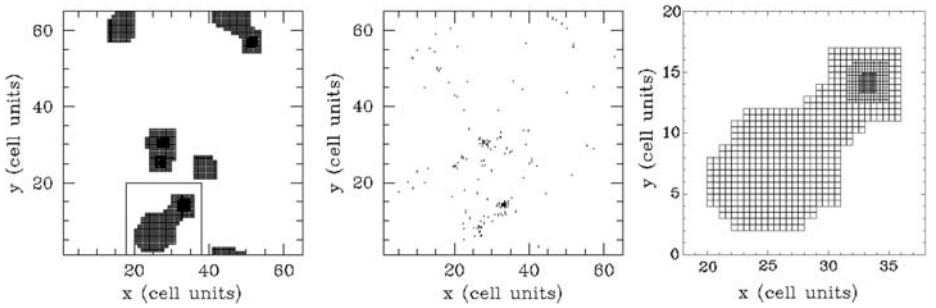


Fig. 2 A slice through the refinement structure (the base grid is not shown) in a ΛCDM simulation (*left panel*) and the corresponding slice through the particle distribution (*middle panel*). The area enclosed by the *square* is enlarged in the *right panel*. Taken from Kravtsov et al. (1997)

where the summation is performed over a cell's neighbours. Here, Δ is the actual spatial resolution of the solution (potential), while $\Delta\tau$ is a fictitious time step (not related to the actual time integration of the N-body system). This finite difference method is stable when $\Delta\tau \leq \Delta^2/6$. More details can be found in Press et al. (1992) and also Kravtsov et al. (1997). Figure 2, from Kravtsov et al. (1997), shows an example of the mesh constructed to calculate the potential in a cosmological simulation.

2.4 Hybrids (*TreePM* / P^3M)

Hybrid methods can be constructed as a synthesis of the particle-mesh method and the tree algorithm. In *TreePM* methods (Xu 1995; Bode et al. 2000; Bagla 2002; Bagla and Ray 2003) the potential is explicitly split in Fourier space into a long-range and a short-range part according to $\Phi_{\mathbf{k}} = \Phi_{\mathbf{k}}^{\text{long}} + \Phi_{\mathbf{k}}^{\text{short}}$, where

$$\Phi_{\mathbf{k}}^{\text{long}} = \Phi_{\mathbf{k}} \exp(-\mathbf{k}^2 r_{\text{s}}^2), \tag{21}$$

with r_{s} describing the spatial scale of the force-split. The long range potential can be computed very efficiently with mesh-based Fourier methods.

The short-range part of the potential can be solved in real space by noting that for $r_{\text{s}} \ll L$ the short-range part of the real-space solution of the Poisson equation is given by

$$\Phi^{\text{short}}(\mathbf{x}) = -G \sum_i \frac{m_i}{\mathbf{r}_i} \text{erfc}\left(\frac{\mathbf{r}_i}{2r_{\text{s}}}\right). \tag{22}$$

Here $\mathbf{r}_i$ is the distance of any particle i to the point $\mathbf{x}$. Thus the short-range force can be computed by the tree algorithm, except that the force law is modified by a long-range cut-off factor.

Such hybrid methods can result in a very substantial improvement of the performance compared to ordinary tree methods. In addition one typically gains accuracy in the long-range force, which is now basically exact, and not an approximation as in the tree method. Furthermore, if r_{s} is chosen to be slightly larger than the mesh scale, force anisotropies, that exist in plain *PM* methods, can be suppressed to essentially arbitrarily low levels. A *TreePM* approach also maintains all the most important advantages of the tree algorithm, namely its

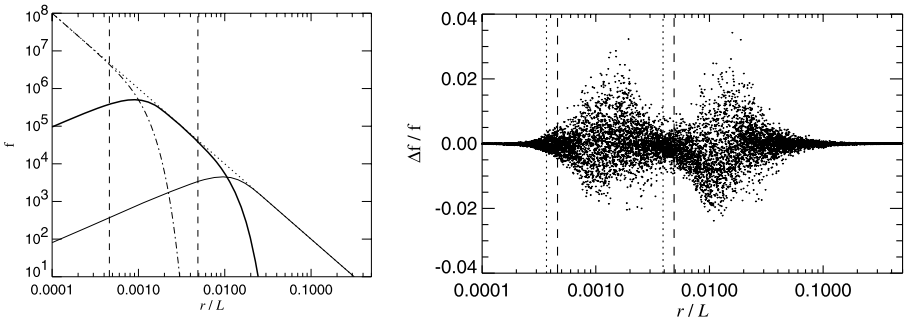


Fig. 3 Force decomposition and force error of the *TreePM* scheme in the case when two stacked meshes are used. The *left panel* illustrates the strength of the short-range (*dot-dashed*), intermediate-range (*thick solid*), and long-range (*solid*) force as a function of distance in a periodic box. The spatial scales of the two splits are marked with *vertical dashed lines*. The *right panel* shows the error distribution of the PM force. The outer matching region exhibits a very similar error characteristic as the inner match of tree- and PM-force. In both cases, for separations of order the fine or coarse mesh scale (*dotted lines*), respectively, force errors of up to 1–2% arise, but the r.m.s. force error stays well below 1%, and the mean force tracks the correct result accurately. Taken from Springel (2005)

insensitivity to clustering, its essentially unlimited dynamical range, and its precise control of the softening scale of the gravitational force.

Figure 3 shows how the force matching works in the *GADGET-2* code (Springel 2005), where such a hybrid method is further extended to two subsequent stacked FFTs combined with the tree algorithm. This extension enables one to increase the dynamic range, which, in turn, improves the computational speed in high resolution simulations of the evolution of galaxy clusters within a large cosmological volume.

Although used much earlier (because much easier to implement), the P^3M method can be seen as a special case of the *TreePM*, where the tree is replaced by the direct sum. Note that also in the tree algorithm the nearest forces are calculated by a direct sum, thus the P^3M approach formally corresponds to extending the direct sum of the tree method to the scale where the *PM* force computation takes over. Couchman (1991) presented an improved version of the P^3M method, by allowing spatially adaptive mesh refinements in regions with high particle density (Adaptive P^3M or AP^3M). The improvement in the performance made it very attractive and several cosmological simulations were performed with this technique, including the Hubble Volume simulations (Evrard et al. 2002).

2.5 Time-Stepping and Integration

The accuracy obtained when evolving the system depends on the size of the time step and on the integrator scheme used. Finding the optimum size of time step is not trivial. A very simple criterion often used is

$$\Delta t = \alpha\sqrt{\epsilon/|\mathbf{a}|}, \tag{23}$$

where $|a|$ is the acceleration obtained at the previous time step, ϵ is a length scale, which can typically be associated with the gravitational softening, and α is a tolerance parameter. More details about different time step criteria can be found for example in Power et al. (2003) and references therein. For the integration of the variables (positions and velocities) of the system, we only need to integrate first order equations of the form $\dot{\mathbf{y}} = f(\mathbf{y})$, e.g. ordinary differential equations (ODEs) with appropriate initial conditions. Note that we can

first solve this ODE for the velocity $\mathbf{v}$ and then treat $\dot{\mathbf{x}} = \mathbf{v}$ as an independent ODE, at basically no extra cost.

One can distinguish implicit and explicit methods for propagating the system from step n to step $n+1$. Implicit methods usually have better properties, however they need to solve the system iteratively, which usually requires inverting a matrix which is only sparsely sampled, and has the dimension of the total number of the data points, namely grid or particle points. Therefore, N-body simulations mostly adopt explicit methods.

The simplest (but never used) method to perform the integration of an ODE is called Euler's method; here the integration is just done by multiplying the derivatives with the length of the time step. The explicit form of such a method can be written as

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{f}(\mathbf{y}_n)\Delta t, \tag{24}$$

whereas the implicit version is written as

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{f}(\mathbf{y}_{n+1})\Delta t. \tag{25}$$

Note that in the latter equation $\mathbf{y}_{n+1}$ appears on the left and right side, which makes it clear why it is called implicit. Obviously the drawback of the explicit method is that it assumes that the derivatives (e.g. the forces) do not change during the time step.

An improvement to this method can be obtained by using the mean derivative during the time step, which can be written with the *implicit mid-point rule* as

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{f}[0.5(\mathbf{y}_n + \mathbf{y}_{n+1})]\Delta t. \tag{26}$$

An explicit rule using the forces at the next time step is the so-called *predictor-corrector method*, where one first predicts the variables for the next time step

$$\mathbf{y}^0_{n+1} = \mathbf{y}_n + \mathbf{f}(\mathbf{y}_n)\Delta t \tag{27}$$

and then uses the forces calculated there to correct this prediction (the so-called corrector step) as

$$\mathbf{y}_{n+1} = \mathbf{y}_n + 0.5[\mathbf{f}(\mathbf{y}_n) + \mathbf{f}(\mathbf{y}^0_{n+1})]\Delta t. \tag{28}$$

This method is accurate to second order.

In fact, all these methods are special cases of the so-called *Runge-Kutta method* (RK), which achieves the accuracy of a Taylor series approach without requiring the calculation of higher order derivatives. The price one has to pay is that the derivatives (e.g. forces) have to be calculated at several points, effectively splitting the interval Δt into special subsets. For example, a second order RK scheme can be constructed by

$$\mathbf{k}_1 = \mathbf{f}(\mathbf{y}_n) \tag{29}$$

$$\mathbf{k}_2 = \mathbf{f}(\mathbf{y}_n + \mathbf{k}_1\Delta t) \tag{30}$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + 0.5(\mathbf{k}_1 + \mathbf{k}_2)\Delta t. \tag{31}$$

In a fourth order RK scheme, the time interval Δt also has to be subsampled to calculate the mid-points, e.g.

$$\mathbf{k}_1 = \mathbf{f}(\mathbf{y}_n, t_n) \tag{32}$$

$$\mathbf{k}_2 = \mathbf{f}(\mathbf{y}_n + \mathbf{k}_1 \Delta t/2, t_n + \Delta t/2) \tag{33}$$

$$\mathbf{k}_3 = \mathbf{f}(\mathbf{y}_n + \mathbf{k}_2 \Delta t/2, t_n + \Delta t/2) \tag{34}$$

$$\mathbf{k}_4 = \mathbf{f}(\mathbf{y}_n + \mathbf{k}_3 \Delta t/2, t_n + \Delta t) \tag{35}$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \left(\frac{\mathbf{k}_1}{6} + \frac{\mathbf{k}_2}{3} + \frac{\mathbf{k}_3}{3} + \frac{\mathbf{k}_4}{6} \right) \Delta t. \tag{36}$$

More details on how to construct the coefficient for an n-th order RK scheme are given in e.g. Chapra and Canale (1997).

Another possibility is to use the so-called *leap-frog method*, where the derivatives (e.g. forces) and the positions are shifted in time by half a time step. This feature can be used to integrate directly the second order ODE of the form $\ddot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. Depending on whether one starts with a *drift* (D) of the system by half a time step or one uses the forces at the actual time to propagate the system (*kick*, K), one obtains a KDK version

$$\mathbf{v}_{n+1/2} = \mathbf{v}_n + \mathbf{f}(\mathbf{x}_n) \Delta t/2 \tag{37}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{v}_{n+1/2} \Delta t \tag{38}$$

$$\mathbf{v}_{n+1} = \mathbf{v}_{n+1/2} + \mathbf{f}(\mathbf{x}_{n+1}) \Delta t/2 \tag{39}$$

or a DKD version of the method

$$\mathbf{x}_{n+1/2} = \mathbf{x}_n + \mathbf{v}_n \Delta t/2 \tag{40}$$

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \mathbf{f}(\mathbf{x}_{n+1/2}) \Delta t \tag{41}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_{n+1/2} + \mathbf{v}_{n+1} \Delta t/2. \tag{42}$$

This method is accurate to second order, and, as will be shown in the next paragraph, also has other advantages. For more details see Springel (2005).

It is also clear that, depending on the application, a lower order scheme applied with more, and thus smaller, time steps can be more efficient than a higher order scheme, which enables the use of larger time steps. In the upper rows of Fig. 4, we show the numerical integration of a Kepler problem (i.e. two point-like masses with large mass difference which orbit around each other like a planet-sun system) of high eccentricity $e = 0.9$, using second-order accurate leap-frog and Runge-Kutta schemes with fixed time step. There is no long-term drift in the orbital energy for the leap-frog result (left panel); only a small residual precession of the elliptical orbit is observed. On the other hand, the second-order Runge-Kutta integrator, which has formally the same error per step, fails catastrophically for an equally large time step (middle panel). After only 50 orbits, the binding energy has increased by $\sim 30\%$. If we instead employ a fourth-order Runge-Kutta scheme using the same time step (right panel), the integration is only marginally more stable, now giving a decline of the binding energy of $\sim 40\%$ over 200 orbits. Note however that such a higher order integration scheme requires several force evaluations per time step, making it computationally much more expensive for a single step than the leap-frog, which requires only one force evaluation per step. The underlying mathematical reason for the remarkable stability of the leap-frog integrator lies in its symplectic properties. For a more detailed discussion, see Springel (2005).

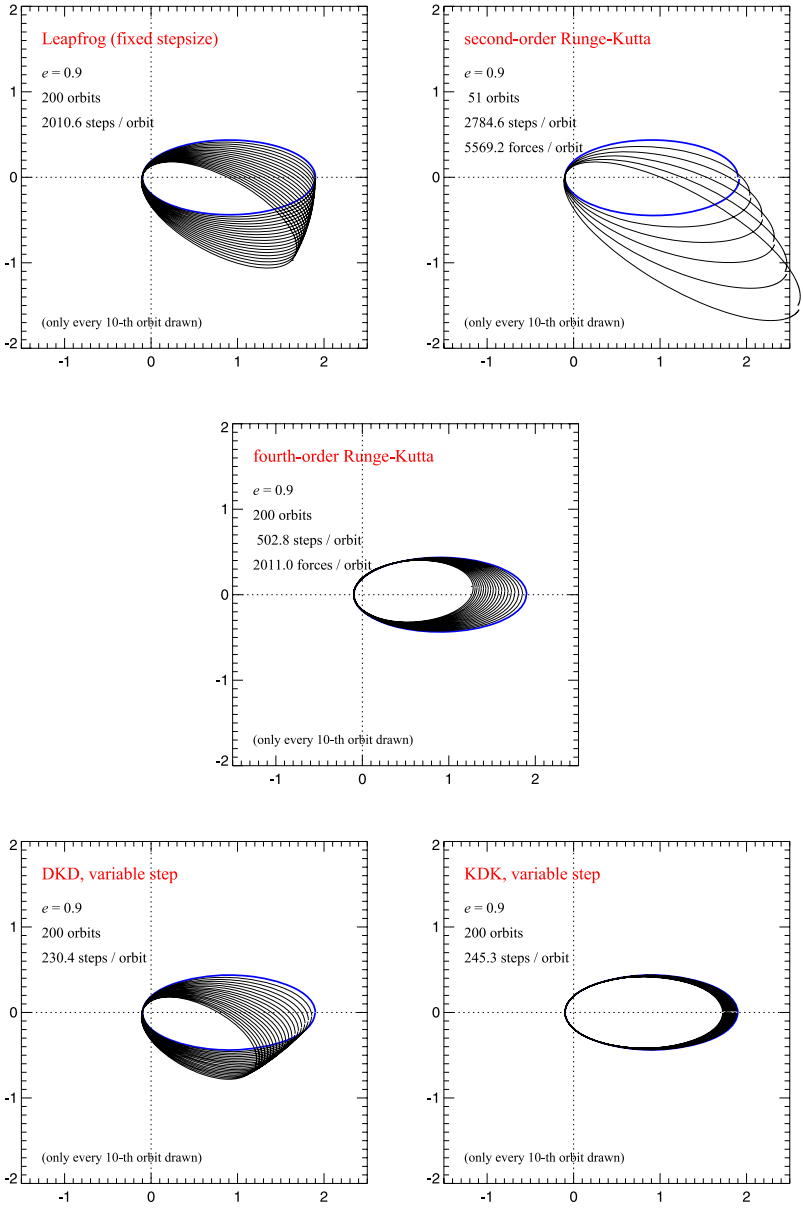


Fig. 4 The *upper two rows* show a Kepler problem of high eccentricity evolved with different simple time integration schemes, using an equal time step in all cases. Even though the leap-frog (*upper left panel*) and the second order Runge-Kutta (*upper right panel*) produce comparable errors in a single step, the long term stability of the integration is very different. Even a computationally much more expensive fourth order Runge-Kutta scheme (*middle row*), with a smaller error per step, performs dramatically worse than the leap-frog in this problem. The lower row shows the same using leap-frog schemes with a variable time step from step to step, based on the $\Delta t \propto 1/\sqrt{|\mathbf{a}|}$ criterion commonly employed in cosmological simulations. As a result of the variable time steps, the integration is no longer manifestly time reversible, and long term secular errors develop. Interestingly, the error in the KDK (Kick-Drift-Kick) variant grows four times more slowly than in the DKD (Drift-Kick-Drift) variant, despite being of equal computational cost. Taken from Springel (2005)

In cosmological simulations, we are confronted with a large dynamic range in timescales. In high-density regions, like at the centres of galaxies, the required time steps are orders of magnitude smaller than in the low-density regions of the intergalactic medium, where a large fraction of the mass resides. Hence, evolving all the particles with the smallest required time step implies a substantial waste of computational resources. An integration scheme with individual time steps tries to cope with this situation more efficiently. The principal idea is to compute forces only for a certain group of particles in a given kick operation (K), with the other particles being evolved on larger time steps being usually just drifted (D) and 'kicked' more rarely.

The KDK scheme is hence clearly superior once one allows for individual time steps, as shown in the lower row of Fig. 4. It is also possible to try to recover the time reversibility more precisely. Hut et al. (1995) discuss an implicit time step criterion that depends both on the beginning and on the end of the time step, and, similarly, Quinn et al. (1997) discuss a binary hierarchy of trial steps that serves a similar purpose. However, these schemes are computationally impractical for large collisionless systems. Fortunately, however, in this case, the danger of building up large errors by systematic accumulation over many periodic orbits is much smaller, because the gravitational potential is highly time-dependent and the particles tend to make comparatively few orbits over a Hubble time.

2.6 Initial Conditions

Having robust and well justified initial conditions is one of the key points of any numerical effort. For cosmological purposes, observations of the large–scale distribution of galaxies and of the CMB agree to good precision with the theoretical expectation that the growth of structures starts from a Gaussian random field of initial density fluctuations; this field is thus completely described by the power spectrum $P(|\mathbf{k}|)$ whose shape is theoretically well motivated and depends on the cosmological parameters and on the nature of Dark Matter.

To generate the initial conditions, one has to generate a set of complex numbers with a randomly distributed phase ϕ and with amplitude normally distributed with a variance given by the desired spectrum (e.g. Bardeen et al. 1986). This can be obtained by drawing two random numbers ϕ in]0, 1] and A in]0, 1] for every point in k-space

$$\hat{\delta}_{\mathbf{k}} = \sqrt{-2P(|\mathbf{k}|)\ln(A)}\mathrm{e}^{i2\pi\phi}. \tag{43}$$

To obtain the perturbation field generated from this distribution, one needs to generate the potential $\Phi(\mathbf{q})$ on a grid $\mathbf{q}$ in real space via a Fourier transform, e.g.

$$\Phi(\mathbf{q}) = \sum_k \frac{\hat{\delta}_{\mathbf{k}}}{\mathbf{k}^2}\mathrm{e}^{i\mathbf{kq}}. \tag{44}$$

The subsequent application of the Zel'dovich approximation (Zel'dovich 1970) enables one to find the initial positions

$$\mathbf{x} = \mathbf{q} - D^+(z)\Phi(\mathbf{q}) \tag{45}$$

and velocities

$$\mathbf{v} = \dot{D}^+(z)\nabla\Phi(\mathbf{q}) \tag{46}$$

of the particles, where $D^+(z)$ and $\dot{D}^+(z)$ indicate the cosmological linear growth factor and its derivative at the initial redshift z. A more detailed description can be found in e.g. Efstathiou et al. (1985).

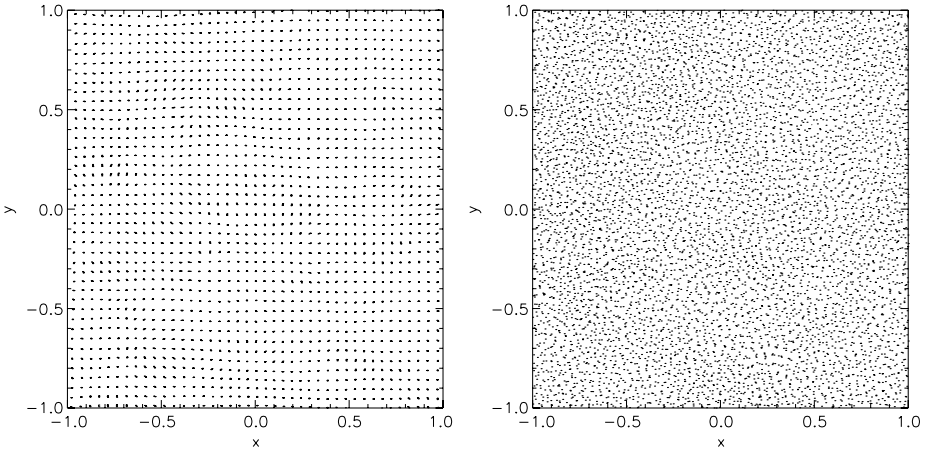


Fig. 5 Shown is a slice to the particle distribution with the imposed displacement, taken from the same cosmological initial conditions, once based on an originally regular grid (*left panel*) and once based on an originally glass like particle distribution (*right panel*)

There are two further complications which should be mentioned. The first is that one can try to reduce the discreteness effect that is induced on the density power spectrum by the regularity of the underlying grid of the particle positions **q** that one has at the start. This can be done by constructing an amorphous, fully relaxed particle distribution to be used, instead of a regular grid. Such a particle distribution can be constructed by applying negative gravity to a system and evolving it for a long time, including a damping of the velocities, until it reaches a relaxed state, as suggested by White (1996). Figure 5 gives a visual impression on the resulting particle distributions.

A second complication is that, even for studying individual objects like galaxy clusters, large-scale tidal forces can be important. A common approach used to deal with this problem is the so-called "zoom" technique: a high resolution region is self-consistently embedded in a larger scale cosmological volume at low resolution (see e.g. Tormen et al. 1997). This approach usually allows an increase of the dynamical range of one to two orders of magnitude while keeping the full cosmological context. For galaxy simulations it is even possible to apply this technique on several levels of refinements to further improve the dynamical range of the simulation (e.g. Stoehr et al. 2003). A frequently used, publicly available package to create initial conditions is the *COSMICS* package by Bertschinger (1995).

2.7 Resolution

There has been a long standing discussion in the literature to understand what is the optimal setup for cosmological simulations, and how many particles are needed to resolve certain regions of interest. Note that the number of particles needed for convergence also depends on what quantity one is interested in. For example, mass functions, which count identified halos, usually give converging results at very small particle numbers per halo ($\approx$30–50), whereas structural properties, like a central density or the virial radius, converge only at significantly higher particle numbers ($\approx$ 1000). As we will see in a later chapter, if one wants to infer hydrodynamical properties like baryon fraction or X-ray luminosity, values converge only for halos represented by even more particles ($\approx$ 10 000).

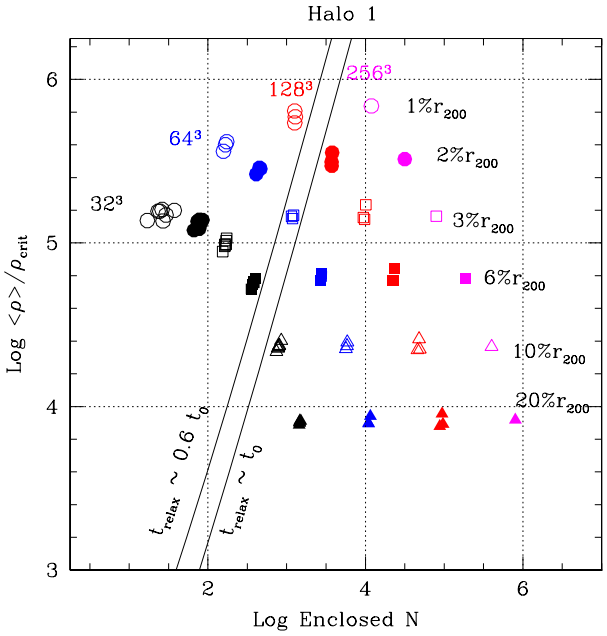


Fig. 6 Mean inner density contrast as a function of the enclosed number of particles in 4 series of simulations varying the number of particles in the high-resolution box, from 32^3 to 256^3. Each *symbol* corresponds to a fixed fraction of the virial radius, as shown by the labels on the right. The number of particles needed to obtain robust results increases with density contrast, roughly as prescribed by the requirement that the collisional relaxation timescale should remain longer than the age of the Universe. According to this, robust numerical estimates of the mass profile of a halo are only possible to the right of the curve labelled $t_{\rm relax} \sim 0.6 t_0$. Taken from Power et al. (2003)

Recently, Power et al. (2003) performed a comprehensive series of convergence tests designed to study the effect of numerical parameters on the structure of simulated CDM halos. These tests explore the influence of the gravitational softening, the time stepping algorithm, the starting redshift, the accuracy of force computations, and the number of particles in the spherically-averaged mass profile of a galaxy-sized halo in the CDM cosmogony with a non-null cosmological constant (ΛCDM). Power et al. (2003), and the references therein, suggest empirical rules that optimise the choice of these parameters. When these choices are dictated by computational limitations, Power et al. (2003) offer simple prescriptions to assess the effective convergence of the mass profile of a simulated halo. One of their main results is summarised in Fig. 6, which shows the convergence of a series of simulations with different mass resolution on different parts of the density profile of a collapsed object. This figure clearly demonstrates that the number of particles within a certain radius needed to obtain converging results depends on the enclosed density.

In general, both the size and the dynamical range or resolution of the simulations have been increasing very rapidly over the last decades. Figure 7 shows a historical compilation of large N-body simulations: their size growth, thanks to improvements in the algorithms, is faster than the underlying growth of the available CPU power.

2.8 Code Comparison for Pure Gravity

In the last thirty years cosmology has turned from a science of order-of-magnitude estimates to a science with accuracies of 10% or less in its measurements and theoretical predictions. Crucial observations along the way were the measurement of the cosmic microwave background radiation, and large galaxy surveys. In the future such observations will yield even higher accuracy (1%) over a much wider dynamical range. Such measurements will provide insight into several topics, e.g. the nature of dark energy (expressed by the equation of state

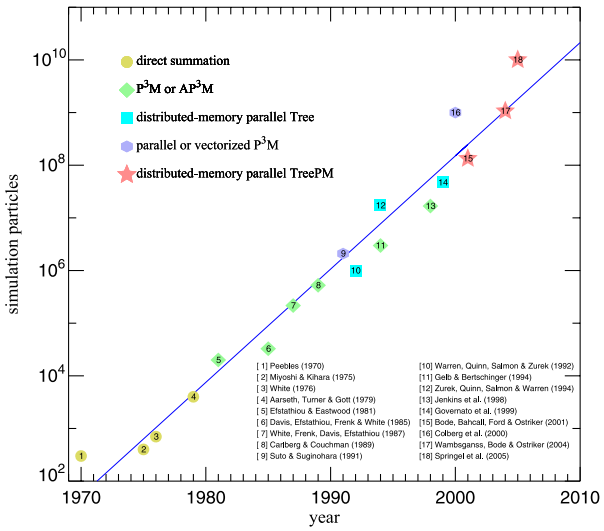


Fig. 7 Moore's empirical law shows that the computing power typically doubles every 18 months. This figure shows the size of N-body simulations as a function of their running date. Clearly, specially recently, the improvement in the algorithms allowed the simulation to grow faster than the improvement of the underlying CPU power. Kindly provided by Volker Springel

$w = p/\rho$ with p being the pressure and ρ the density). In order to make optimal use of the observations, theoretical calculations of at least the same level of accuracy are required. As physics in the highly non-linear regime, combined with complicated gas physics and astrophysical feedback processes are involved, this represents a real challenge.

Different numerical methods have therefore to be checked and compared continuously. The most recent comparison of ten commonly-used codes from the literature has been performed in an extensive comparison program. The ten codes used for the comparison performed by Heitmann et al. (2007) cover a variety of methods and application arenas. The simulation methods employed include parallel particle-in-cell (*PIC*) techniques (the *PM* codes MC^2 and *PMM*, the Particle-Mesh/Adaptive Mesh Refinement (*AMR*) codes *ENZO* and *FLASH*), a hybrid of *PIC* and direct N-body (the AP^3M code *Hydra*), tree algorithms (the treecodes *PKDGRAV* and *HOT*), and hybrid tree-*PM* algorithms (*GADGET-2*, *TPM*, and *TreePM*).

The results from the code comparisons are satisfactory and not unexpected, but also show that much more work is needed in order to attain the required accuracy for upcoming surveys. The halo mass function is a very stable statistic, the agreement over wide ranges of mass being better than 5%. Additionally, the low mass cutoff for individual codes can be reliably predicted by a simple criterion.

The internal structure of halos in the outer regions of $\sim R_{200}$ also appears to be very similar between different simulation codes. Larger differences between the codes in the inner region of the halos occur if the halo is not in a relaxed state: in this case, time stepping issues might also play an important role (e.g. particle orbit phase errors, global time mismatches). For halos with a clear single centre, the agreement is very good and predictions for the fall-off of the profiles from resolution criteria hold as expected. The investigation of the halo counts as a function of density revealed an interesting problem with the *TPM* code, the simulation suffering from a large deficit in medium density regimes. The *AMR* codes showed a large deficit of small halos over almost the entire density regime, as the base grid of the *AMR* simulation sets a resolution limit that is too low for the halos, as can be seen in Fig. 8.

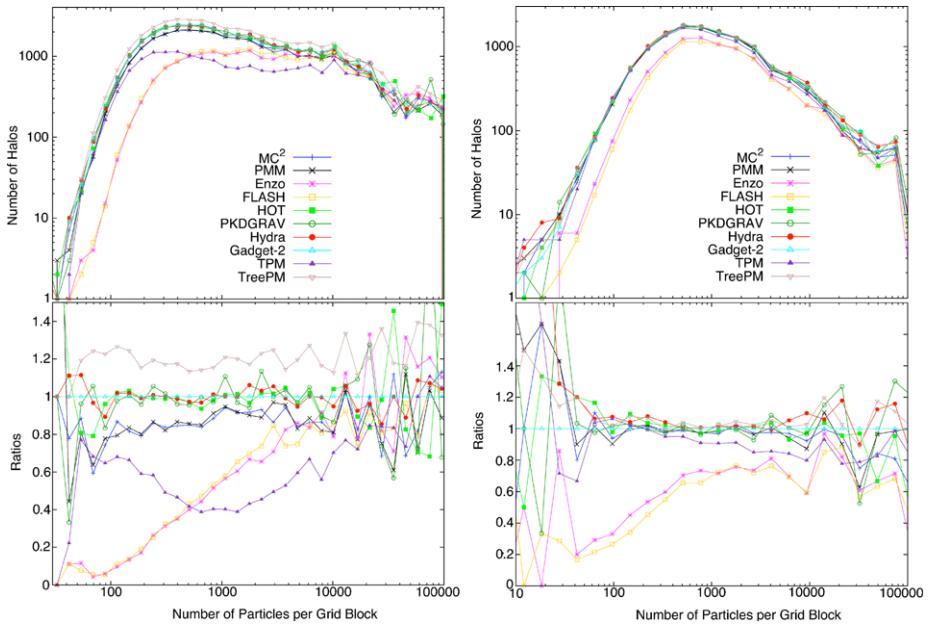


Fig. 8 A recent comparison of the predicted number of halos as a function of density for ten different cosmological codes. *Left panel*: halos with 10–40 particles, *right panel*: halos with 41–2500 particles. The *lower panels* show the residuals with respect to *GADGET-2*. Both panels show the deficit of small halos in *ENZO* and *FLASH* over most of the density region—only at very high densities do the results catch up. The behaviour of the *TPM* simulation is interesting: not only does this simulation have a deficit of small halos but the deficit is very significant in medium density regions, in fact falling below the two Adaptive Mesh Refinement codes. The slight excess of small halos shown in the *TreePM* run vanishes completely if the halo cut is raised to 20 particles per halo and the *TreePM* results are in that case in excellent agreement with *GADGET-2*. Adapted from Heitmann et al. (2007)

The power spectrum measurements revealed definitively more scatter among the different codes than expected. The agreement in the nonlinear regime is at the 5–10% level, even on moderate spatial scales around $k = 10h$ Mpc^{-1}. This disagreement on small scales is connected to differences of the codes in the inner regions of the halos. For more detailed discussion see Heitmann et al. (2007) and references therein.

In a detailed comparison of *ENZO* and *GADGET*, O'Shea et al. (2005) already pointed out that to reach reasonable good agreement, relatively conservative criteria for the adaptive grid refinement are needed. Furthermore, choosing a grid resolution twice as high as the mean inter-particle distance of the dark matter particles is recommended, to improve the small scale accuracy of the calculation of the gravitational forces.

3 Hydro Methods

The baryonic content of the Universe can typically be described as an ideal fluid. Therefore, to follow the evolution of the fluid, one usually has to solve the set of hydrodynamic equations

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} - \nabla\Phi, \tag{47}$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho \nabla \mathbf{v} = 0 \tag{48}$$

and

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho} \nabla \cdot \mathbf{v} - \frac{\Lambda(u, \rho)}{\rho}, \tag{49}$$

which are the *Euler equation*, *continuity equation* and the *first law of thermodynamics*, respectively. They are closed by an *equation of state*, relating the pressure P to the internal energy (per unit mass) u. Assuming an ideal, monatomic gas, this will be

$$P = (\gamma - 1)\rho u \tag{50}$$

with $\gamma = 5/3$. In the next sections, we will discuss how to solve this set of equations, neglecting radiative losses described by the cooling function $\Lambda(u, \rho)$; in Sect. 4.1 we will give examples of how radiative losses or additional sources of heat are included in cosmological codes. We can also assume that the $\nabla\Phi$ term will be solved using the methods described in the previous section.

As a result of the high nonlinearity of gravitational clustering in the Universe, there are two significant features emerging in cosmological hydrodynamic flows; these features pose more challenges than the typical hydrodynamic simulation without self-gravity. One significant feature is the extremely supersonic motion around the density peaks developed by gravitational instability, which leads to strong shock discontinuities within complex smooth structures. Another feature is the appearance of an enormous dynamic range in space and time, as well as in the related gas quantities. For instance, the hierarchical structures in the galaxy distribution span a wide range of length scales, from the few kiloparsecs resolved in an individual galaxy to the several tens of megaparsecs characterising the largest coherent scale in the Universe.

A variety of numerical schemes for solving the coupled system of collisional baryonic matter and collisionless dark matter have been developed in the past decades. They fall into two categories: particle methods, which discretise mass, and grid-based methods, which discretise space. We will briefly describe both methods in the next two sections.

3.1 Eulerian (Grid)

The set of hydrodynamical equations for an expanding Universe reads

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v} \cdot \nabla)\mathbf{v} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a\rho}\nabla P - \frac{1}{a}\nabla\Phi, \tag{51}$$

$$\frac{\partial \rho}{\partial t} + \frac{3\dot{a}}{a}\rho + \frac{1}{a}\nabla \cdot (\rho \mathbf{v}) = 0 \tag{52}$$

and

$$\frac{\partial}{\partial t}(\rho u) + \frac{1}{a}\mathbf{v} \cdot \nabla(\rho u) = -(\rho u + P)\left(\frac{1}{a}\nabla \cdot \mathbf{v} + 3\frac{\dot{a}}{a}\right) \tag{53}$$

respectively, where the right term in the last equation reflects the expansion in addition to the usual $P\mathrm{d}V$ work.

The grid-based methods solve these equations based on structured or unstructured grids, representing the fluid. One distinguishes *primitive variables*, which determine the thermodynamic properties, (e.g. ρ, $\mathbf{v}$ or P) and *conservative variables* which define the conservation

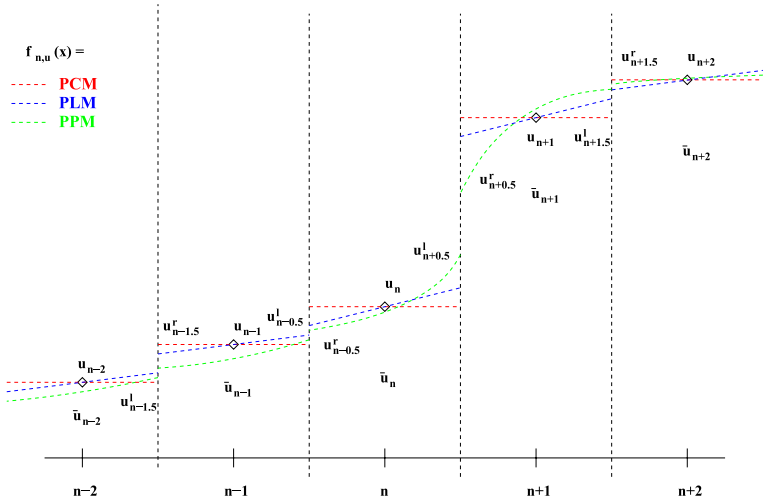


Fig. 9 Reconstruction of the principal variables (u_n) on the grid using different methods, like *piecewise constant* (PCM), *piecewise linear* (PLM) or *piecewise parabolic* (PPM). The reconstruction scheme then allows one to calculate cell averages ($\bar{u}_n$) as well as the left and right-hand sided values on the cell boundaries ($u^l_{n\pm0.5}, u^r_{n\pm0.5}$)

laws, (e.g. ρ, $\rho\mathbf{v}$ or ρu). Early attempts were made using a central difference scheme, where fluid is only represented by the centred cell values (e.g. *central variables*, u_n in Fig. 9 and derivatives are obtained by the finite-difference representation, similar to (15) and (16), see for example Cen (1992). Such methods will however break down in regimes where discontinuities appear. These methods therefore use artificial viscosity to handle shocks (similar to the smoothed particle hydrodynamics method described in the next section). Also, by construction, they are only first-order accurate.

More modern approaches use reconstruction schemes, which, depending on their order, take several neighbouring cells into account to reconstruct the field of any hydrodynamical variable. Figure 9 illustrates three different reconstruction schemes, with increasing order of accuracy, as *piecewise constant method* (*PCM*), *piecewise linear method* (*PLM*, e.g. Colella and Glaz 1985) and *piecewise parabolic method* (*PPM*, Colella and Woodward 1984). The shape of the reconstruction $f_{n,u}(x)$ is then used to calculate the total integral of a quantity over the grid cell, divided by the volume of each cell (e.g. cell average, $\hat{u}_n$), rather than pointwise approximations at the grid centres (e.g. central variables, u_n).

$$\hat{u}_n = \int_{x_{n-0.5}}^{x_{n+0.5}} f_{n,u}(x)\mathrm{d}x. \tag{54}$$

They are also used to calculate the left and right-hand sided values at the cell boundaries (e.g. $u^l_{n\pm0.5}, u^r_{n\pm0.5}$), which are used later as initial conditions to solve the Riemann problem. To avoid oscillations (e.g. the development of new extrema), additional constraints are included in the reconstruction. For example, in the *PLM* reconstruction this is ensured by using so-called slope limiters which estimate the maximum slope allowed for the reconstruction. One way is to demand that the total variation among the interfaces does not increase with time. Such so-called total variation diminishing schemes (*TVD*, Harten 1983), nowadays provide various different slope limiters suggested by different authors. In our example

illustrated by Fig. 9, the so called minmod slope limiter

$$\overline{\Delta u_i} = \text{minmod}\left(\Theta(u_{i+1} - u_i), (u_{i+1} - u_{i-1})/2, \theta(u_i - u_{i-1})\right), \tag{55}$$

where $\overline{\Delta u_i}$ is the limiter slope within the cell i and $\theta = [1, 2]$, would try to fix the slope $f'_{n-1,u}(x_{n-1})$ and $f'_{n,u}(x_n)$, such as to avoid that $u^l_{n-0.5}$ becomes larger than $u^r_{n-0.5}$. The so called Aldaba-type limiter

$$\overline{\Delta u_i} = \frac{2(u_{i+1} - u_i)(u_i - u_{i-1}) + \epsilon}{(u_{i+1} - u_i)^2 + (u_i - u_{i-1})^2 + \epsilon^2} \frac{1}{2}(u_{i+1} - u_{i-1}), \tag{56}$$

where ϵ is a small positive number to avoid problems in homogeneous regions, would try to avoid that $u^l_{n-0.5}$ is getting larger than u_n and that $u^r_{n-0.5}$ is getting smaller than u_{n-1}, e.g. that a monotonic profile in u_i is preserved.

In the *PPM* (or even higher order) reconstruction this enters as an additional condition when finding the best-fitting polynomial function. The additional cells which are involved in the reconstruction are often called the *stencil*. Modern, high order schemes usually have stencils based on at least 5 grid points and implement essentially non-oscillatory (*ENO*; Harten et al. 1987) or monoticity preserving (*MP*) methods for reconstruction, which maintain high-order accuracy. For every reconstruction, a *smoothness indicator* S^m_n can be constructed, which is defined as the integral over the sum of the squared derivatives of the reconstruction over the stencil chosen, e.g.

$$S^m_n = \sum_{l=1}^{2} \int_{x_{n-m}}^{x_{n+m}} (\Delta x)^{2l-1} \left(\partial^l_x f^m_{n,u}(x)\right)^2 \mathrm{d}V. \tag{57}$$

In the *ENO* schemes, a set of candidate polynomials p^m_n with order $2m + 1$ for a set of *stencils* based on different numbers of grid cells m are used to define several different reconstruction functions $f^m_{n,u}$. Then, the reconstruction with the lowest smoothness indicator S^m_n is chosen. In this way the order of reconstruction will be reduced around discontinuities, and oscillating behaviour will be suppressed.

To improve on the *ENO* schemes in robustness and accuracy one can, instead of selecting the reconstruction with the best smoothness indicator S^m_n, construct the final reconstruction by building the weighted reconstruction

$$\hat{f}_u(x) = \sum_m w_m f^m_{n,u}(x), \tag{58}$$

where the weights w_m are a proper function of the smoothness indicators S^m_n. This procedure is not unique. Jiang and Shu (1996) proposed defining

$$w_m = \frac{\alpha_m}{\sum_l \alpha_l} \tag{59}$$

with

$$\alpha_l = \frac{C_l}{(\epsilon + S^l_n)^\beta}, \tag{60}$$

where C_l, ϵ and β are free parameters, which for example can be taken from Levy et al. (1999). This are the so-called weighted essentially non-oscillatory (*WENO*) schemes. These

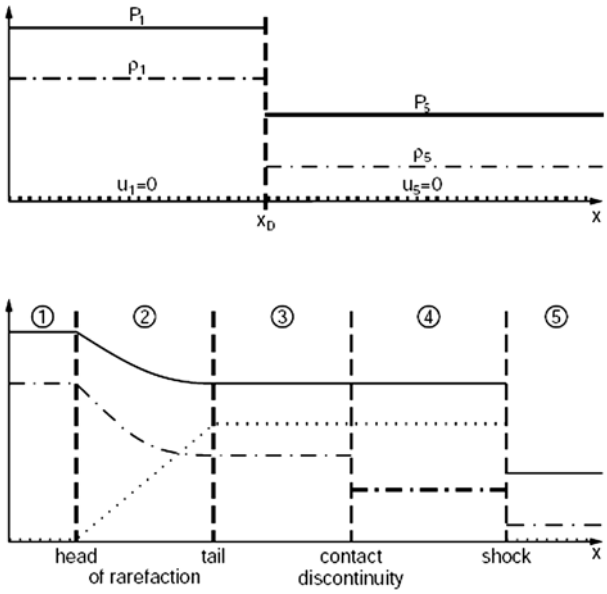


Fig. 10 The Riemann problem: The *upper panel* shows the initial state, the *lower panel* shows the evolved problem for the case of no relative motion between the two sides ($u_1 = u_5 = 0$). The *solid lines* mark the pressure P, the *dashed dotted lines* the density ρ and the *dotted line* the velocity v. Kindly provided by Ewald Müller

schemes can simultaneously provide a high-order resolution for the smooth part of the solution and a sharp, monotonic shock or contact discontinuity transition. For a review on *ENO* and *WENO* schemes, see e.g. Shu (1998).

After the left and right-hand values at the cell boundaries (e.g. interfaces) are reconstructed, the resulting Riemann problem is solved, e.g. the evolution of two constant states separated by a discontinuity. This can be done either analytically or approximately, using left and right-handed values at the interfaces as a jump condition.

With the solution one obtains, the fluxes across these boundaries for the time step can be calculated and the cell averages $\hat{u}_n$ can be updated accordingly. In multiple dimensions, all these steps are performed for each coordinate direction separately, taking only the interface values along the individual axes into account. There are attempts to extend the reconstruction schemes, to directly reconstruct the principal axis of the Riemann problem in multiple dimensions, so that then it has to be solved only once for each cell. However the complexity of reconstructing the surface of shocks in three dimensions has so far seen to be untraceable.

How to solve the general Riemann problem, e.g. the evolution of a discontinuity initially separating two states, can be found in text books (e.g. Courant and Friedrichs 1948). Here we give only the evolution of a shock tube as an example. This corresponds to a system where both sides are initially at rest. Figure 10 shows the initial and the evolved system. The latter can be divided into 5 regions. The values for regions 1 and 5 are identical to the initial configuration. Region 2 is a rarefaction wave which is determined by the states in regions 1 and 3. Therefore we are left with 6 variables to determine, namely ρ_3, P_3, v_3 and ρ_4, P_4, v_4, where we have already eliminated the internal energy u_i in all regions, as it can be calculated from the equation of state. As there is no mass flux through the contact discontinuity, and as the pressure is continuous across the contact discontinuity, we can eliminate two of the six variables by setting $v_3 = v_4 = v_c$ and $P_3 = P_4 = P_c$. The general *Rankine-Hugoniot* conditions, describing the jump conditions at a discontinuity, read

$$\rho_l v_l = \rho_r v_r \tag{61}$$

$$\rho_l v_l^2 + P_l = \rho_r v_r^2 + P_r \tag{62}$$

$$v_l \left(\rho_l (v_l^2/2 + u_l) + P_l\right) = v_r \left(\rho_r (v_r^2/2 + u_r) + P_r\right), \tag{63}$$

where we have assumed a coordinate system which moves with the shock velocity v_s. Assuming that the system is at rest in the beginning, e.g. $v_1 = v_5 = 0$, the *first* Rankine-Hugoniot condition for the shock between region 4 and 5 moving with a velocity v_s (note the implied change of the coordinate system) is in our case

$$m = \rho_5 v_\mathrm{s} = \rho_4 (v_\mathrm{s} - v_\mathrm{c}) \tag{64}$$

and therefore the shock velocity becomes

$$v_\mathrm{s} = \frac{\rho_4 v_\mathrm{c}}{\rho_4 - \rho_5}. \tag{65}$$

The *second* Rankine-Hugoniot condition is

$$m v_\mathrm{c} = \rho_4 (v_\mathrm{s} - v_\mathrm{c}) v_\mathrm{c} = P_\mathrm{c} - P_5, \tag{66}$$

which, combined with the first, can be written as

$$\rho_4 \left(\frac{\rho_4 v_\mathrm{c}}{\rho_4 - \rho_5} - v_\mathrm{c}\right) v_\mathrm{c} = P_\mathrm{c} - P_5, \tag{67}$$

which, slightly simplified, leads to a *first condition*

$$(P_5 - P_\mathrm{c}) \left(\frac{1}{\rho_5} - \frac{1}{\rho_4}\right) = -v_\mathrm{c}^2. \tag{68}$$

The *third* Rankine-Hugoniot condition is

$$m \left(\epsilon_4 + \frac{v_\mathrm{c}^2}{2} - \epsilon_5\right) = P_\mathrm{c} v_\mathrm{c}, \tag{69}$$

which, by eliminating m, can be written as

$$\epsilon_4 - \epsilon_5 = \frac{P_\mathrm{c} + P_5}{P_\mathrm{c} - P_5} \frac{v_\mathrm{c}^2}{2}. \tag{70}$$

Using the first condition (68) and assuming an ideal gas for the equation of state, one gets

$$\frac{1}{\gamma - 1} \left(\frac{P_\mathrm{c}}{\rho_4} - \frac{P_5}{\rho_5}\right) = \frac{P_\mathrm{c} + P_5}{2\rho_4 \rho_5}, \tag{71}$$

which leads to the *second condition*

$$\frac{P_\mathrm{c} - P_5}{P_\mathrm{c} + P_5} = \gamma \frac{\rho_4 - \rho_5}{\rho_4 + \rho_5}. \tag{72}$$

The *third condition* comes from the fact that the entropy ($\propto \ln(P/\rho^\gamma)$ stays constant in the rarefaction wave, and therefore one can write it as

$$\frac{P_1}{P_\mathrm{c}} = \left(\frac{\rho_1}{\rho_3}\right)^\gamma. \tag{73}$$

The *fourth condition* comes from the fact that the Riemann Invariant

$$v + \int \frac{c}{\rho} \mathrm{d}\rho \tag{74}$$

is a constant, which means that

$$v_\mathrm{c} + \int \frac{c_3}{\rho_3} \mathrm{d}\rho = v_1 + \int \frac{c_1}{\rho_1} \mathrm{d}\rho, \tag{75}$$

where $c = \sqrt{\gamma P/\rho}$ denotes the sound velocity, with which the integral can be written as

$$\int \frac{c}{\rho} \mathrm{d}\rho = \frac{2}{\gamma - 1} \sqrt{\frac{\gamma P}{\rho}}. \tag{76}$$

Therefore, the fourth condition can be written as

$$v_\mathrm{c} + \frac{2}{\gamma - 1} \sqrt{\frac{\gamma P_\mathrm{c}}{\rho_3}} = \frac{2}{\gamma - 1} \sqrt{\frac{\gamma P_1}{\rho_1}}. \tag{77}$$

Combining all 4 conditions (68, 72, 73 and 77) and defining the initial density ratio $\lambda = \rho_1/\rho_5$ one gets the non linear, algebraic equation

$$\frac{\rho_1}{\rho_5} \frac{1}{\lambda} \frac{(1-P)^2}{\gamma(1+P) - 1 + P} = \frac{2\gamma}{(\gamma-1)^2} \left[1 - \left(\frac{P}{\lambda} \right)^{(\gamma-1)/(2\gamma)} \right]^2 \tag{78}$$

for the pressure ratio $P = P_\mathrm{c}/P_5$. Once P_c is known from solving this equation, the remaining unknowns can be inferred step by step from the four conditions.

There are various approximate methods to solve the Riemann problem, including the so-called *ROE method* (e.g. Powell et al. 1999), *HLL/HLLE* method (e.g. see Harten et al. 1983; Einfeldt 1988; Einfeldt et al. 1991) and *HLLC* (e.g. see Li 2005). A description of all these methods is outside the scope of this review, so we redirect the reader to the references given or textbooks like LeVeque (2002).

At the end of each time step, one has to compute the updated central values u_n from the updated cell average values $\hat{u}_n$. Normally, this would imply inverting (54), which is not trivial in the general case. Therefore, usually an additional constraint is placed on the reconstruction method, namely that the reconstruction fulfills $u_n = \hat{u}_n$. In this case the last step is trivial.

In general, the grid-based methods suffer from limited spatial resolution, but they work extremely well in both low- and high-density regions, as well as in shocks. In cosmological simulations, accretion flows with large Mach numbers (e.g. $M > 100$) are very common. Here, following the total energy in the hydrodynamical equations, one can get inaccurate thermal energy, leading to negative pressure, due to discretisation errors when the kinetic energy dominates the total energy. In such cases, as suggested by Ryu et al. (1993) and Bryan et al. (1995), the numerical schemes usually switch from formulations solving the total energy to formulations based on solving the internal energy in these hypersonic flow regions.

In the cosmological setting, there are the *TVD*-based codes including those of Ryu et al. (1993) and Li et al. (2006) (*CosmoMHD*), the moving-mesh scheme (Pen 1998) and the *PLM*-based code *ART* (Kravtsov et al. 1997; Kravtsov 2002). The *PPM*-based codes include

those of Stone and Norman (1992) (*Zeus*), Bryan et al. (1995) (*ENZO*), Ricker et al. (2000) (*COSMOS*) and Fryxell et al. (2000) (*FLASH*). There is also the *WENO*-based code by Feng et al. (2004).

3.2 Lagrangian (SPH)

The particle methods include variants of smoothed particle hydrodynamics (*SPH*; Gingold and Monaghan 1977; Lucy 1977) such as those of Evrard (1988), Hernquist and Katz (1989), Navarro and White (1993), Couchman et al. (1995) (*Hydra*), Steinmetz (1996a) (*GRAPE-SPH*), Owen et al. (1998), and Springel et al. (2001a), Springel (2005) (*GADGET*). The *SPH* method solves the Lagrangian form of the Euler equations and can achieve good spatial resolutions in high-density regions, but it works poorly in low-density regions. It also suffers from degraded resolution in shocked regions due to the introduction of a sizable artificial viscosity. Agertz et al. (2007) argued that whilst Eulerian grid-based methods are able to resolve and treat dynamical instabilities, such as *Kelvin-Helmholtz* or *Rayleigh-Taylor*, these processes are poorly resolved by existing *SPH* techniques. The reason for this is that *SPH*, at least in its standard implementation, introduces spurious pressure forces on particles in regions where there are steep density gradients, in particular near contact discontinuities. This results in a boundary gap of the size of an *SPH* smoothing kernel radius, over which interactions are severely damped. Nevertheless, in the cosmological context, the adaptive nature of the *SPH* method compensates for such shortcomings, thus making *SPH* the most commonly used method in numerical hydrodynamical cosmology.

3.2.1 Basics of SPH

The basic idea of *SPH* is to discretise the fluid by mass elements (e.g. particles), rather than by volume elements as in the Eulerian methods. Therefore it is immediately clear that the mean inter-particle distance in collapsed objects will be smaller than in underdense regions; the scheme will thus be adaptive in spatial resolution by keeping the mass resolution fixed. For a comprehensive review see Monaghan (1992). To build continuous fluid quantities, one starts with a general definition of a kernel smoothing method

$$\langle A(\mathbf{x})\rangle = \int W(\mathbf{x}-\mathbf{x}', h) A(\mathbf{x}') \mathrm{d}\mathbf{x}', \tag{79}$$

which requires that the kernel is normalised (i.e. $\int W(\mathbf{x}, h)\mathrm{d}\mathbf{x} = 1$) and collapses to a delta function if the smoothing length h approaches zero, namely $W(\mathbf{x}, h) \rightarrow \delta(\mathbf{x})$ for $h \rightarrow 0$.

One can write down the continuous fluid quantities (e.g. $\langle A(\mathbf{x})\rangle$) based on the discretised values A_j represented by the set of the individual particles m_j at the position $\mathbf{x}_j$ as

$$\langle A_i\rangle = \langle A(\mathbf{x}_i)\rangle = \sum_j \frac{m_j}{\rho_j} A_j W(\mathbf{x}_i - \mathbf{x}_j, h)\,, \tag{80}$$

where we assume that the kernel depends only on the distance modulus (i.e. $W(|\mathbf{x}-\mathbf{x}'|, h)$) and we replace the volume element of the integration, $\mathrm{d}\mathbf{x} = \mathrm{d}^3x$, with the ratio of the mass and density m_j/ρ_j of the particles. Although this equation holds for any position $\mathbf{x}$ in space, here we are only interested in the fluid representation at the original particle positions $\mathbf{x}_i$, which are the only locations where we will need the fluid representation later on. It is important to note that for kernels with compact support (i.e. $W(\mathbf{x}, h) = 0$ for $|\mathbf{x}| > h$) the summation does not have to be done over all the particles, but only over the particles within

the sphere of radius h, namely the neighbours around the particle i under consideration. Traditionally, the most frequently used kernel is the B_2-spline, which can be written as

$$W(x,h) = \frac{\sigma}{h^\nu} \begin{cases} 1 - 6\left(\frac{x}{h}\right)^2 + 6\left(\frac{x}{h}\right)^3, & 0 \le \frac{x}{h} < 0.5, \\ 2\left(1 - \frac{x}{h}\right)^3, & 0.5 \le \frac{x}{h} < 1, \\ 0, & 1 \le \frac{x}{h}, \end{cases} \tag{81}$$

where ν is the dimensionality (e.g. 1, 2 or 3) and σ is the normalisation

$$\sigma = \begin{cases} \frac{16}{3}, & \nu = 1, \\ \frac{80}{7\pi}, & \nu = 2, \\ \frac{8}{\pi}, & \nu = 3. \end{cases} \tag{82}$$

Sometimes, spline kernels of higher order are used for very special applications; however the B_2 spline kernel turns out to be the optimal choice in most cases.

When one identifies A_i with the density ρ_i, ρ_i cancels out on the right hand side of (80), and we are left with the density estimate

$$\langle \rho_i \rangle = \sum_j m_j W(\mathbf{x}_i - \mathbf{x}_j, h), \tag{83}$$

which we can interpret as the density of the fluid element represented by the particle i.

Now even derivatives can be calculated as

$$\nabla \langle A_i \rangle = \sum_j \frac{m_j}{\rho_j} A_j \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h), \tag{84}$$

where ∇_i denotes the derivative with respect to $\mathbf{x}_i$. A pairwise symmetric formulation of derivatives in *SPH* can be obtained by making use of the identity

$$(\rho \nabla) \cdot A = \nabla(\rho \cdot A) - \rho \cdot (\nabla A), \tag{85}$$

which allows one to re-write a derivative as

$$\nabla \langle A_i \rangle = \frac{1}{\rho_i} \sum_j m_j (A_j - A_i) \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h). \tag{86}$$

Another way of symmetrising the derivative is to use the identity

$$\frac{\nabla A}{\rho} = \nabla\left(\frac{A}{\rho}\right) + \frac{A}{\rho^2} \nabla \rho, \tag{87}$$

which then leads to the following form of the derivative:

$$\nabla \langle A_i \rangle = \rho_i \sum_j m_j \left(\frac{A_j}{\rho_j^2} + \frac{A_i}{\rho_i^2}\right) \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h). \tag{88}$$

3.2.2 The Fluid Equations

By making use of these identities, the Euler equation can be written as

$$\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = -\sum_j m_j \left(\frac{P_j}{\rho_j^2} + \frac{P_i}{\rho_i^2} + \Pi_{ij} \right) \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h). \tag{89}$$

By combining the above identities and averaging the result, the term $-(P/\rho)\nabla \cdot \mathbf{v}$ from the first law of thermodynamics can similarly be written as

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \frac{1}{2}\sum_j m_j \left(\frac{P_j}{\rho_j^2} + \frac{P_i}{\rho_i^2} + \Pi_{ij} \right) \left(\mathbf{v}_j - \mathbf{v}_i\right) \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h). \tag{90}$$

Here we have added a term Π_{ij} which is the so-called *artificial viscosity*. This term is usually needed to capture shocks and its construction is similar to other hydro-dynamical schemes. Usually, one adopts the form proposed by Monaghan and Gingold (1983) and Balsara (1995), which includes a bulk viscosity and a von Neumann-Richtmeyer viscosity term, supplemented by a term controlling angular momentum transport in the presence of shear flows at low particle numbers (Steinmetz 1996b). Modern schemes implement a form of the artificial viscosity as proposed by Monaghan (1997) based on an analogy with Riemann solutions of compressible gas dynamics. To reduce this artificial viscosity, at least in those parts of the flows where there are no shocks, one can follow the idea proposed by Morris and Monaghan (1997): every particle carries its own artificial viscosity, which eventually decays outside the regions which undergo shocks. A detailed study of the implications on the ICM of such an implementation can be found in Dolag et al. (2005).

The *continuity equation* does not have to be evolved explicitly, as it is automatically fulfilled in Lagrangian methods. As shown earlier, density is no longer a variable but can be, at any point, calculated from the particle positions. Obviously, mass conservation is guaranteed, unlike volume conservation: in other words, the sum of the volume elements associated with all of the particles might vary with time, especially when strong density gradients are present.

3.2.3 Variable Smoothing Length

Usually, the smoothing length h will be allowed to vary for each individual particle i and is determined by finding the radius h_i of a sphere which contains n neighbours. Typically, different numbers n of neighbours are chosen by different authors, ranging from 32 to 80. In principle, depending on the kernel, there is an optimal choice of neighbours (e.g. see Silverman 1986 or similar books). However, one has to find a compromise between a large number of neighbours, leading to larger systematics but lower noise in the density estimates (especially in regions with large density gradients) and a small number of neighbours, leading to larger sample variances for the density estimation. In general, once every particle has its own smoothing length, a symmetric kernel $W(\mathbf{x}_i - \mathbf{x}_j, h_i, h_j) = \bar{W}_{ij}$ has to be constructed to keep the conservative form of the formulations of the hydrodynamical equations. There are two main variants used in the literature: one is the kernel average $\bar{W}_{ij} = (W(\mathbf{x}_i - \mathbf{x}_j, h_i) + W(\mathbf{x}_i - \mathbf{x}_j, h_j))/2$, the other is an average of the smoothing length $\bar{W}_{ij} = W(\mathbf{x}_i - \mathbf{x}_j, (h_i + h_j)/2)$. The former is the most commonly used approach.

Note that in all of the derivatives discussed above, it is assumed that h does not depend on the position $\mathbf{x}_j$. Thus, by allowing the smoothing length h_i to be variable for each

particle, one formally neglects the correction term $\partial W/\partial h$, which would appear in all the derivatives. In general, this correction term cannot be computed trivially and therefore many implementations do not take it into account. It is well known that such formulations are poor at conserving numerically both internal energy and entropy at the same time, independently of the use of internal energy or entropy in the formulation of the first law of thermodynamics, see Hernquist (1993). In the next subsection, we present a way of deriving the equations which include these correction terms $\partial W/\partial h$; this equation set represents a formulation which conserves numerically both entropy and internal energy.

3.2.4 The Entropy Conservation Formalism

To derive a better formulation of the *SPH* method, Springel and Hernquist (2002) started from the entropic function $A = P/\rho^\gamma$, which will be conserved in adiabatic flows. The internal energy per unit mass can be inferred from this entropic function as

$$u_i = \frac{A_i}{\gamma - 1}\rho_i^{\gamma-1} \tag{91}$$

at any time, if needed. Entropy will be generated by shocks, which are captured by the artificial viscosity Π_{ij} and therefore the entropic function will evolve as

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = \frac{1}{2}\frac{\gamma-1}{\rho_i^{\gamma-1}}\sum_j m_j \Pi_{ij}\left(\mathbf{v}_j - \mathbf{v}_i\right)\nabla_i \bar{W}_{ij}. \tag{92}$$

The Euler equation can be derived starting by defining the Lagrangian of the fluid as

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\sum_i m_i \dot{\mathbf{x}}_i^2 - \frac{1}{\gamma-1}\sum_i m_i A_i \rho_i^{\gamma-1} \tag{93}$$

which represents the entire fluid and has the coordinates $\mathbf{q} = (\mathbf{x}_1, \ldots, \mathbf{x}_N, h_1, \ldots, h_N)$. The next important step is to define constraints, which allow an unambiguous association of h_i for a chosen number of neighbours n. This can be done by requiring that the kernel volume contains a constant mass for the estimated density,

$$\phi_i(\mathbf{q}) = \frac{4\pi}{3}h_i^3\rho_i - n m_i = 0. \tag{94}$$

The equation of motion can be obtained as the solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\mathbf{q}}_i} - \frac{\partial L}{\partial \mathbf{q}_i} = \sum_j \lambda_j \frac{\partial \phi_j}{\partial \mathbf{q}_i}, \tag{95}$$

which—as demonstrated by Springel and Hernquist (2002)—can be written as

$$\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = -\sum_j m_j \left(f_j \frac{P_j}{\rho_j^2}\nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h_j) + f_i \frac{P_i}{\rho_i^2}\nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h_i) + \Pi_{ij}\nabla_i \bar{W}_{ij}\right), \tag{96}$$

where we already have included the additional term due to the artificial viscosity Π_{ij}, which is needed to capture shocks. The coefficients f_i incorporate fully the variable smoothing

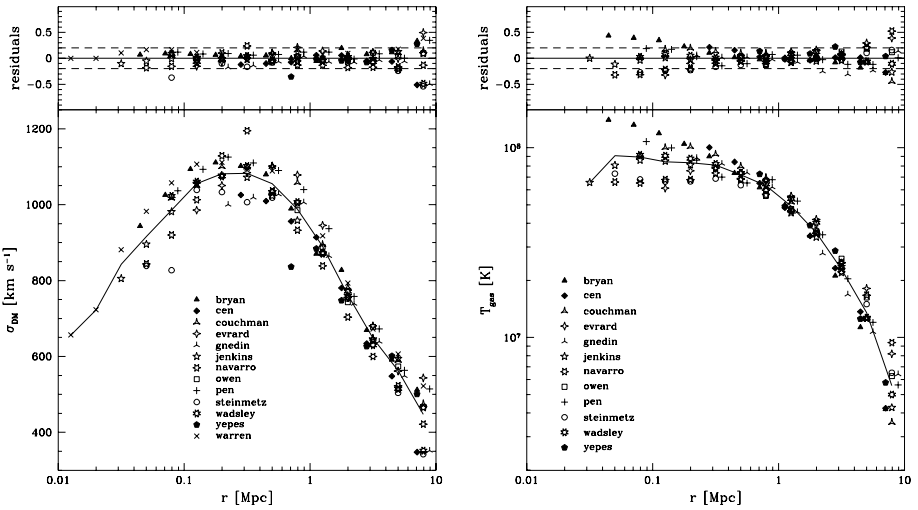


Fig. 11 One-dimensional velocity dispersion profile (*left panel*) and gas temperature profile (*right panel*) of the cluster at $z = 0$ of the Santa Barbara Comparison Project (Frenk et al. 1999). The *solid line* is the profile averaged over the 12 simulations. The *symbols* correspond to individual simulations. The *crosses* in the *left panel* correspond to a dark-matter only simulation. The *top panels* show the residual from the mean profile. Taken from Frenk et al. (1999)

length correction term and are defined as

$$f_i = \left(1 + \frac{h_i}{3\rho_i}\frac{\partial \rho_i}{\partial h_i}\right)^{-1}. \tag{97}$$

Note that in addition to the correction terms, which can be easily calculated together with the density estimation, this formalism also avoids all the ambiguities we saw in the derivations of the equations in the previous section. This formalism defines how the kernel averages (symmetrisation) have to be taken, and also fixes h_i to unambiguous values. For a detailed derivation of this formalism and its conserving capabilities see Springel and Hernquist (2002).

3.3 Code Comparison

The Eulerian and Lagrangian approaches described in the previous sections should provide the same results when applied to the same problem, like the interaction of multi-phase fluids (Agertz et al. 2007). To verify that the code correctly solves the hydrodynamical set of equations, each code is usually tested against problems whose solution is known analytically. In practise, these are shock tubes or spherical collapse problems. In cosmology, a relevant test is to compare the results provided by the codes when they simulate the formation of cosmic structure, when finding an analytic solution is impractical; for example O'Shea et al. (2005) compare the thermodynamical properties of the intergalactic medium predicted by the *GADGET* (*SPH*-based) and *ENZO* (grid-based) codes. Another example of a comparison between grid-based and *SPH*-based codes can be found in Kang et al. (1994).

A detailed comparison of hydrodynamical codes which simulate the formation and evolution of a cluster was provided by the Santa Barbara Cluster Comparison Project (Frenk

et al. 1999). Frenk et al. (1999) comprised 12 different groups, each using a code either based on the *SPH* technique (7 groups) or on the grid technique (5 groups). Each simulation started with identical initial conditions of an individual massive cluster in a flat CDM model with zero cosmological constant. Each group was free to decide resolution, boundary conditions and the other free parameters of their code. The simulations were performed ignoring radiative losses and the simulated clusters were compared at $z = 0.5$ and $z = 0$.

The resulting dark matter properties were similar: it was found a 20% scatter around the mean density and velocity dispersion profiles (left panel of Fig. 11). A similar agreement was also obtained for many of the gas properties, like the temperature profile (right panel of Fig. 11) or the ratio of the specific dark matter kinetic energy and the gas thermal energy.

Somewhat larger differences are present for the inner part of the temperature or entropy profiles and more recent implementations have not yet cured this problem. The largest discrepancy was in the total X-ray luminosity. This quantity is proportional to the square of the gas density, and resolving the cluster central region within the core radius is crucial: the simulations resolving this region had a spread of 2.6 in the total X-ray luminosity, compared to a spread of 10 when all the simulations were included. Frenk et al. (1999) also concluded that a large fraction of the discrepancy, when excluding the X-ray luminosity result, was due to differences in the internal timing of the simulations: these differences produce artificial time shifts between the outputs of the various simulations even if the outputs are formally at the same cosmic time. This reflects mainly the underlying dark matter treatment, including chosen force accuracy, different integration schemes and choice of time steps used, as described in the previous sections. A more worrisome difference between the different codes is the predicted baryon fraction and its profile within the cluster. Here modern schemes still show differences (e.g. see Ettori et al. 2006; Kravtsov et al. 2005), which makes it difficult to use simulations to calibrate the systematics in the cosmological test based on the cluster baryon fraction.

To date, the comparisons described in the literature show a satisfactory agreement between the two approaches, with residual discrepancies originating from the known weaknesses which are specific to each scheme. A further limitation of these comparisons is that, in most cases, the simulations are non-radiative. However, at the current state of the art, performing comparisons of simulations including radiative losses is not expected to provide robust results. As described in the next section, the first relevant process that needs to be added is radiative cooling: however, depending on the square of the gas density, cooling increases with resolution without any indication of convergence, see for example Fig. 13, taken from Borgani et al. (2006). At the next level of complexity, star formation and supernova feedback occur in regions which have a size many orders of magnitude smaller than the spatial resolution of the cosmological simulations. Thus, simulations use phenomenological recipes to describe these processes, and any comparison would largely test the agreement between these recipes rather than identify the inadequacy of the numerical integration schemes.

4 Adding Complexity

In this section, we will give a brief overview of how astrophysical processes, that go beyond the description of the gravitational instability and of the hydrodynamical flows are usually included in simulation codes.

4.1 Cooling

We discuss here how the $\Lambda(u, \rho)$ term is usually added in the first law of thermodynamics, described by (49), and its consequences.

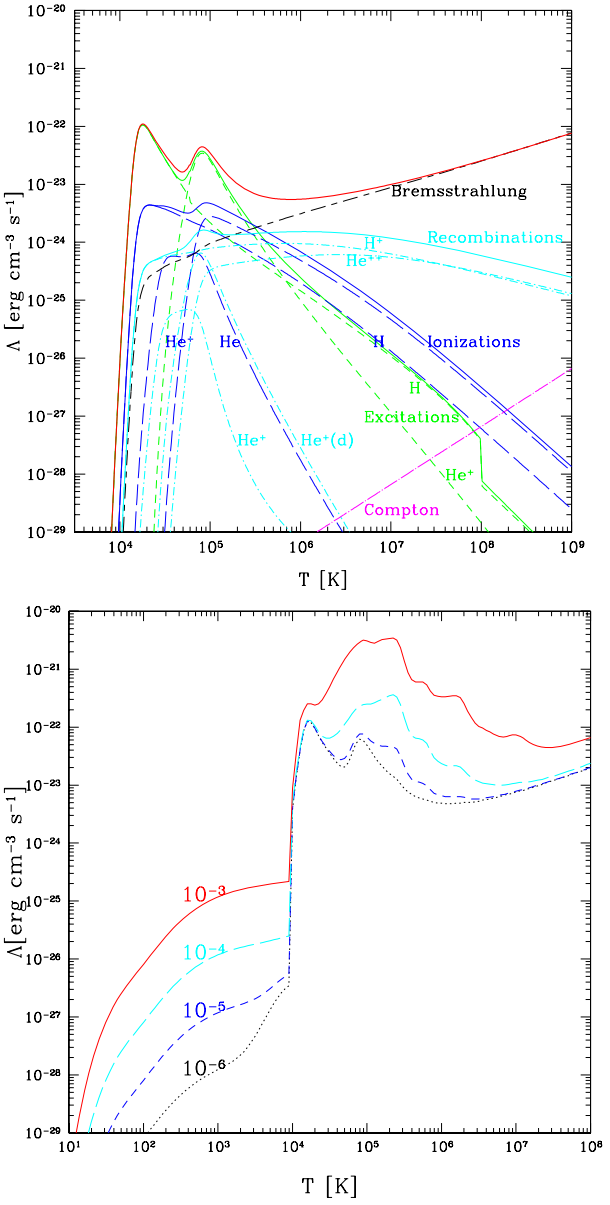


Fig. 12 The *top panel* shows the total cooling curve (*solid line*) and its composition from different processes for a primordial mixture of H and He. The *bottom panel* shows how the total cooling curve will change as a function of different metallicity, as indicated in the plot (in absolute values). The part below 10^4 K also takes into account cooling by molecules (e.g. HD and H_2) and metal lines. Taken from Maio et al. (2007)

In cosmological applications, one is usually interested in structures with virial temperatures larger than 10^4 K. In standard implementations of the cooling function $\Lambda(u, \rho)$, one assumes that the gas is optically thin and in ionisation equilibrium. It is also usually assumed that three-body cooling processes are unimportant, so as to restrict the treatment to two-body processes. For a plasma with primordial composition of H and He, these processes are collisional excitation of H I and He II, collisional ionisation of H I, He I and He II, standard recombination of H II, He II and He III, dielectric recombination of He II, and free-free emission (Bremsstrahlung). The collisional ionisation and recombination rates depend only

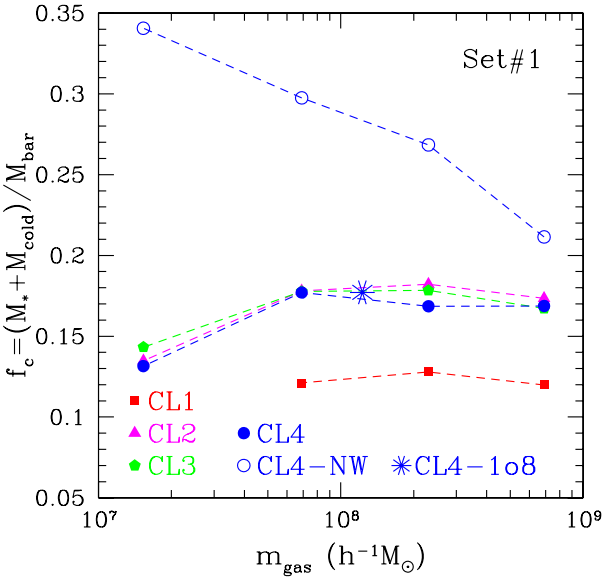


Fig. 13 The fraction of cooled baryons f_c as a function of the mass of the gas particle, for 4 different clusters at different resolutions is shown. *Filled symbols* are for the runs including kinetic feedback (e.g. winds), the *open circles* are re-simulations of one of the clusters with wind feedback turned off. The *asterisk* is for one of the clusters run at very high resolution using fewer, but 8 times heavier, gas particles than normal, so that the gas particle mass is similar to that of the DM particles in the high-resolution region. Taken from Borgani et al. (2006)

on temperature. Therefore, in the absence of ionising background radiation one can solve the resulting rate equation analytically. This leads to a cooling function $\Lambda(u)/\rho^2$ as illustrated in the top panel of Fig. 12. In the presence of ionising background radiation, the rate equations can be solved iteratively. Note that for a typical cosmological radiation background (e.g. UV background from quasars, see Haardt and Madau 1996), the shape of the cooling function can be significantly altered, especially at low densities. For a more detailed discussion see for example Katz et al. (1996). Additionally, the presence of metals will drastically increase the possible processes by which the gas can cool. As it becomes computationally very demanding to calculate the cooling function in this case, one usually resorts to a pre-computed, tabulated cooling function. As an example, the bottom panel of Fig. 12, at temperatures above 10^5 K, shows the tabulated cooling function by Sutherland and Dopita (1993) for different metallicities of the gas, keeping the ratios of the different metal species fixed to solar values. Note that almost all implementations solve the above rate equations (and therefore the cooling of the gas) as a "sub time step" problem, decoupled from the hydrodynamical treatment. In practise this means that one assumes that the density is fixed across the time step. Furthermore, the time step of the underlying hydrodynamical simulation are in general, for practical reasons, not controlled by or related to the cooling time-scale. The resulting uncertainties introduced by these approximations have not yet been deeply explored and clearly leave room for future investigations.

For the formation of the first objects in halos with virial temperatures below 10^4 K, the assumption of ionisation equilibrium no longer holds. In this case, one has to follow the non-equilibrium reactions, solving the balance equations for the individual levels of each species during the cosmological evolution. In the absence of metals, the main coolants are H_2 and H_2^+ molecules (see Abel et al. 1997). HD molecules can also play a significant role. When metals are present, many more reactions are available and some of these can contribute significantly to the cooling function below 10^4 K. This effect is clearly visible in the bottom panel of Fig. 12, for $T < 10^4$ K. For more details see Galli and Palla (1998) or Maio et al. (2007) and references therein.

4.2 Star Formation and Feedback

Including radiative losses in simulations causes two numerical problems. Firstly, cooling is a runaway process and, at the typical densities reached at the centres of galaxy clusters, the cooling time becomes significantly shorter than the Hubble time. As a consequence, a large fraction of the baryonic component can cool down and condense out of the hot phase. Secondly, since cooling is proportional to the square of the gas density, its efficiency is quite sensitive to the presence of the first collapsing small halos, where cooling takes place, and therefore on numerical resolution.

To deal with these issues, one has to include in the code a suitable recipe to convert the reservoir of cold and dense gas into collisionless stars. Furthermore, this stellar component should represent the energy feedback from supernova explosions, which ideally would heat the cold gas, so as to counteract the cooling catastrophe.

As for star formation, a relatively simple recipe is that originally introduced by (Katz et al. 1996), which is often used in cosmological simulations. According to this prescription, for a gas particle to be eligible to form stars, it must have a convergent flow,

$$\nabla \mathbf{v}_i < 0\,, \tag{98}$$

and have density in excess of some threshold value, e.g.

$$\rho_i > 0.1 \text{ atoms}\,\text{cm}^{-3}. \tag{99}$$

These criteria are complemented by requiring the gas to be Jeans unstable, that is

$$\frac{h_i}{c_i} > \frac{1}{\sqrt{4\pi \mathrm{G}\rho_i}}\,, \tag{100}$$

where h_i is either the *SPH* smoothing length or the mesh size for Eulerian codes and c_i is the local sound speed. This indicates that the individual resolution element gets gravitationally unstable. At high redshift, the physical density can easily exceed the threshold given in (99), even for particles not belonging to virialised halos. Therefore one usually applies a further condition on the gas overdensity,

$$\frac{\rho_i}{\rho_{\text{mean}}} > 55.7, \tag{101}$$

which restricts star formation to collapsed, virialised regions. Note that the density criterion is the most important one. Particles fulfilling it in almost all cases also fulfill the other two criteria.

Once a gas particle is eligible to form stars, its star formation rate can be written as

$$\frac{\mathrm{d}\rho_*}{\mathrm{d}t} = -\frac{\mathrm{d}\rho_i}{\mathrm{d}t} = \frac{c_*\rho_i}{t_*}\,, \tag{102}$$

where c_* is a dimensionless star formation rate parameter and t_* the characteristic timescale for star formation. The value of this timescale is usually taken to be the maximum of the dynamical time $t_{\text{dyn}} = (4\pi \mathrm{G}\rho_i)^{-0.5}$ and the cooling time $t_{\text{cool}} = u_i/(\mathrm{d}u_i/\mathrm{d}t)$. In principle, to follow star formation, one would like to produce continuously collisionless star particles. However, for computational and numerical reasons, one approximates this process by waiting for a significant fraction of the gas particle mass to have formed stars according to the above rate; when this is accomplished, a new, collisionless "star" particle is created

from the parent star-forming gas particle, whose mass is reduced accordingly. This process takes place until the gas particle is entirely transformed into stars. In order to avoid spurious numerical effects, which arise from the gravitational interaction of particles with widely differing masses, one usually restricts the number of star particles (so called *generations*) spawned by a gas particle to be relatively small, typically 2–3. Note that it is also common to restrict the described star-formation algorithm to only convert a gas particle into a star particle, which correspond to the choice of only one *generation*. In this case star and gas particles have always the same mass.

To get a more continuous distribution of star particle masses, the probability of forming a star can be written as

$$p = 1 - \exp\left(-c_* \frac{\Delta t}{t_g}\right) \tag{103}$$

and a random number is used to decide when to form a star particle.

According to this scheme of star formation, each star particle can be identified with a Simple Stellar Population (SSP), i.e. a coeval population of stars characterised by a given assumed initial mass function (IMF). Further, assuming that all stars with masses larger than 8 $M_\odot$ will end as type-II supernovae (SN II), one can calculate the total amount of energy (typically 10^{51} erg per supernova) that each star particle can release to the surrounding gas. Under the approximation that the typical lifetime of massive stars which explode as SN II does not exceed the typical time step of the simulation, this is done in the so-called "instantaneous recycling approximation", with the feedback energy deposited in the surrounding gas in the same step.

Improvements with respect to this model include an explicit sub-resolution description of the multi-phase nature of the interstellar medium, which provides the reservoir of star formation. Such a sub grid model tries to model the global dynamical behaviour of the interstellar medium in which cold, star-forming clouds are embedded in a hot medium.

One example is the multi-phase sub-grid model suggested by Springel and Hernquist (2003), in which a star-forming gas particle has a multi-phase nature, with a cold phase describing cold clouds embedded in pressure equilibrium within a hot medium. According to this model, star formation takes place in a self-regulated way. Within this model (as within other models), part of the feedback energy is channelled back into kinetic energy, effectively leading to a quasi self-consistent modelling of galactic outflows, driven by the star-forming regions. Once an efficient form of kinetic feedback is included, the amount of stars formed in simulations turns out to converge when resolution is varied. An example of the predicted stellar component in different galaxy clusters for varying spatial resolutions, taken from Borgani et al. (2006) is shown in Fig. 13.

A further direction for improvement is provided by a more accurate description of stellar evolution, and of the chemical enrichment associated with star forming particles. More accurate models require that energy feedback and metals are released not just by SN II, but also by SN Ia and low and intermediate mass stars, thereby avoiding the instantaneous recycling approximation (e.g., see Borgani et al. 2008—Chap. 18, this issue, for a more detailed discussion of this point).

4.3 Additional Physics

A number of other physical processes, besides those related to star formation and SN feedback, are in general expected to play a role in the evolution of the cosmic baryons and, as such, should be added into the treatment of the hydrodynamical equations. For instance,

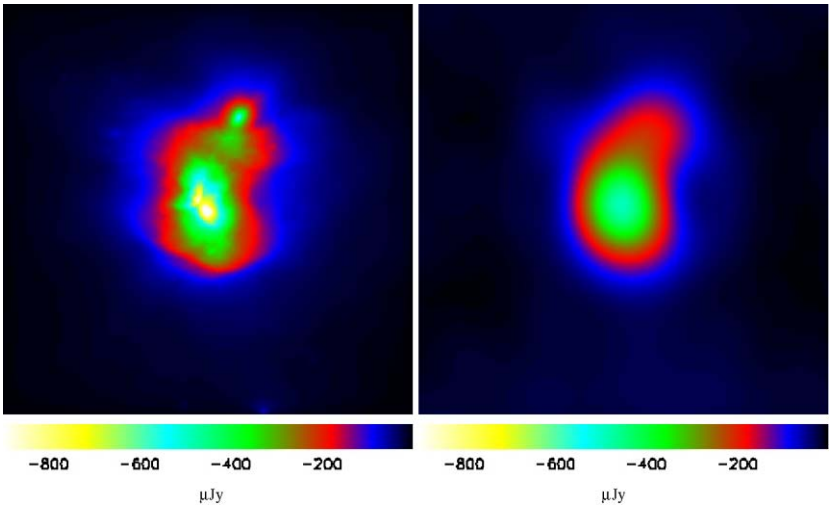


Fig. 14 Maps for the SZ decrement for a simulated galaxy cluster. The original map extracted from the hydrodynamic simulation, and the same map in the simulated observation ($t = 34$ h) which assumes the *AMI* interferometric response, are shown in the *left* and *right panel*, respectively. The side of each map corresponds to 16 arcmin. The colour scale is shown at the *bottom* of each panel. Taken from Bonaldi et al. (2007)

magnetic fields or the effect of a population of relativistic particles as described in detail in the review by Dolag et al. (2008)—Chap. 12, this issue. Other processes, whose effect has been studied in cosmological simulations of galaxy clusters include thermal conduction (Jubelgas et al. 2004; Dolag et al. 2004), radiative transfer to describe the propagation of photons in a medium (e.g. Iliev et al. 2006 and references therein), growth of black holes and the resulting feedback associated with the extracted energy (e.g. Sijacki et al. 2007 and references therein). A description of these processes is far outside the scope of this review and we point the interested reader to the original papers cited above.

5 Connecting Simulations to Observations

Both the recent generation of instruments and an even more sophisticated next generation of instruments (at various wavelengths) will allow us to study galaxy clusters in rich detail. Therefore, when comparing observations with simulations, instrumental effects like resolution and noise, as well as more subtle effects within the observational processes, have to be taken into account to separate true features from instrumental effects and biases. Therefore, the building up of synthetic instruments to "observe" simulations becomes more and more important.

As an example, Fig. 14 shows the difference between a Sunyaev-Zel'dovich map obtained from a simulated galaxy cluster, and how it would be observed with the AMI[1] instrument, assuming the response according to the array configuration of the radio dishes and adding the appropriate noise. Here a 34 hours observation has been assumed, and the *CLEAN* deconvolution algorithm applied. For details see Bonaldi et al. (2007).

[1] http://www.mrao.cam.ac.uk/telescope/ami

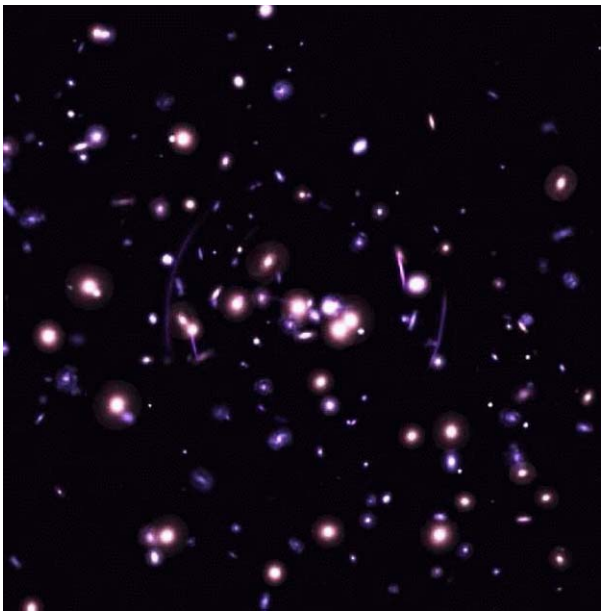


Fig. 15 Composite $ugr + riz + izy$ image of a simulated galaxy cluster including its simulated lensing signal, having its imprint in several strong lensing features. To construct the surface brightness distributions of both the cluster members and of the background galaxies, shapelet decompositions were used. For each cluster galaxy the morphological classifications and spectral energy distribution inferred from the semi-analytic modelling based on the merger trees from the underlying cluster simulation were used to realistically model the optical appearance of the cluster. The field of view is $100'' \times 100''$ and an exposure of 1000 s for each band was assumed. Taken from Meneghetti et al. (2007)

Another example of a synthetic observation by a virtual telescope is shown in Fig. 15, which shows an optical image of a simulated galaxy cluster with several strong lensing features. Here, Meneghetti et al. (2007) investigated the capability of the planned *DUNE* mission to measure the properties of gravitational arcs, including instrumental effects as well as the disturbance by light from the cluster galaxies. Shapelet decompositions—based on galaxy images retrieved from the *GOODS-ACS* archive—were used to describe the surface brightness distributions of both the cluster members and the background galaxies. For the cluster members the morphological classifications inferred from the semi-analytic modelling, based on the merger tree of the underlying cluster simulations were used to assign a spectral energy distribution to each. Several sources of noise like photon noise, sky background, seeing, and instrumental noise are included. For more details see Meneghetti et al. (2007).

Thanks to its high temperature (around 10^8 K) and its high density (up to 10^{-3} particles per cm^3) the intra-cluster medium is an ideal target to be studied by X-ray telescopes. Therefore, so far, most effort towards understanding systematic in the observational process of galaxy clusters has been spent interpreting X-ray observations.

For a direct comparison of the simulations with X-ray observations one has to calculate from the simulated physical quantities (density, temperature, ...) the observed quantities (e.g. surface brightness). The other way—going from observations to physical quantities—is much more difficult and would require a number of assumptions. Fortunately, the ICM is usually optically thin, so that absorption of photons within the ICM does not have to be taken into account. Hence to obtain an X-ray image of the modelled cluster, one has to choose a projection direction and integrate over all the emission of the elements along the line of sight for each pixel in the image. The X-ray emission at each element is usually approximated as the product of the square of the density and the cooling function. As X-ray detectors are only sensitive in a certain energy range, one must be careful to take the correct range. For very realistic images one needs to taken into account also the effects of the X-ray telescopes and detectors (e.g. the limited resolution of the X-ray telescope (point spread function) or

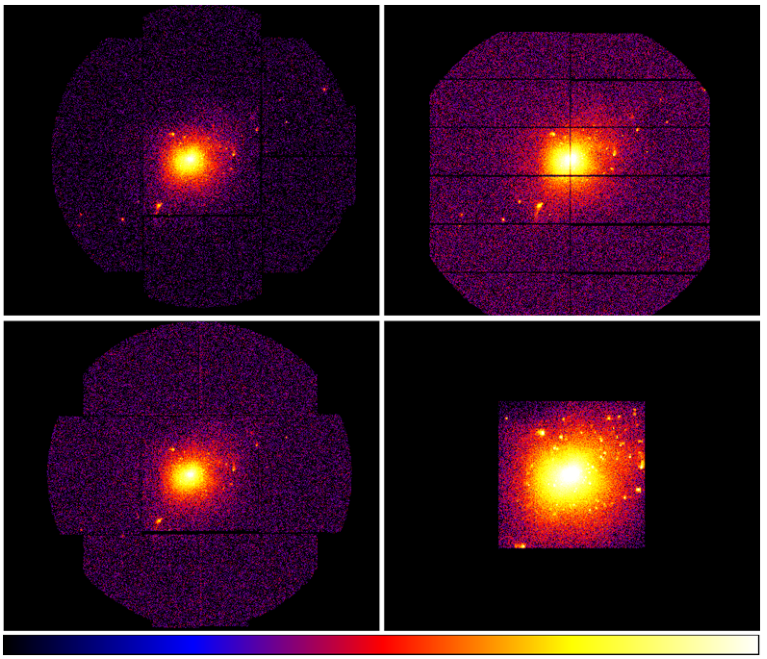


Fig. 16 Simulated photon images in the 0.7–2 keV energy band of a simulated galaxy cluster using *XMAS-2*. The images are binned to 3.2 arcsec. They include background, vignetting effects, out-of-time events and the telescope optical paths. From *top left* to *bottom right* are simulations for the MOS1, PN, and MOS2 instruments on board of the *XMM-Newton* satellite and for the ACIS-S3 instrument on board of the *Chandra* satellite. Kindly provided by Elena Rasia, see Rasia et al. (2006, 2007)

the energy and angle dependent sensitivity (detector response matrix and vignetting). As observed images also depend on the distance of the cluster and the (frequency dependent) absorption of the X-ray emission in the Galaxy these should also be accounted for. For an exact prediction, the emitted spectrum of each element has to be taken into account, as well as many of the effects mentioned above are energy dependent.

Even more difficult is the comparison of quantities derived from spectral analyses such as ICM temperature and ICM metallicity, or their projections in profiles and maps. A pixel in an observed temperature map, for example, is derived using all the photons within the pixel area, and the derived spectrum is fitted with a single model of a hot plasma (see Kaastra et al. 2008—Chap. 9, this issue). These photons, however, come from different positions along the line of sight, that have different emissivities, different temperature and different metallicities. So the spectrum is actually a composite of many different spectra. Obviously such a multi-temperature spectrum cannot be fitted very well by a single temperature model. Sometimes temperature maps are produced in an even cruder way using the ratio of two or more energy bands. These are called hardness ratio maps.

To calculate temperatures, temperature profiles or temperature maps are calculated using mostly emission weighted temperatures, i.e. summing the temperature of all elements along the line of sight, weighted by their emission

$$T = \frac{\int WT\mathrm{d}V}{\int W\mathrm{d}V} \tag{104}$$

with T being the gas temperature and W a weighting factor. Usually W is proportional to the emissivity of each gas element,

$$W = \Lambda(T)n^2 \tag{105}$$

where $\Lambda(T)$ is the cooling curve and n the gas density.

This simple procedure does not of course take into account the shape of the spectra corresponding to gas of different temperatures. It therefore gives only a rough estimate of the actual temperature. With numerical simulations it was investigated how accurate these emission weighted temperatures are, by comparing them with temperatures obtained by actually adding spectra, so-called spectroscopic temperatures. It was found that the emission-weighted temperatures are systematically higher than the spectroscopic temperatures (Mathiesen and Evrard 2001; Gardini et al. 2004; Mazzotta et al. 2004). To overcome this problem, Mazzotta et al. (2004) suggested an approximation to the spectroscopic temperature, the "spectroscopic-like" temperature $T_{\rm sl}$

$$T_{\rm sl} = \frac{\int n^2 T^\alpha / T^{1/2} {\rm d}V}{\int n^2 T^\alpha / T^{3/2} {\rm d}V} \tag{106}$$

which yields for $\alpha = 0.75$ a good estimate of the spectroscopic temperature. In addition the simulated M–T relation is strongly affected, if the emission-weighted temperature is used (Rasia et al. 2005).

The inhomogeneous temperature and metal distribution was also found to be responsible also for inaccurate metallicity measurements. Rasia et al. (2007) studied with numerical simulations, together with the programme *X-MAS2*, how well the elements Fe, O, Mg and Si can be measured in clusters of different temperature. They find that Fe and Si are generally measurable with good accuracy, while O and Mg can be considerably overestimated in hot clusters. Using simulations and the programme *SPEX* (see Kaastra et al. 2008—Chap. 9, this issue) Kapferer et al. (2007) found that due to the metal inhomogeneities the metal mass in clusters is systematically underestimated—in extreme cases by up to a factor of three.

An example of synthetic X-ray observations is shown in Fig. 16. Here *XMAS-2* is used to produce photon images from a simulated galaxy cluster. Shown are simulations for the MOS1, PN, and MOS2 instruments on board the *XMM-Newton* satellite, and for the ACIS-S3 instrument on board the *Chandra* satellite. They include instrumental effects like background, vignetting and response for the individual instruments. For more details see Rasia et al. (2006) and Rasia et al. (2007).

6 Outlook

In the future, the demand on precision in both simulation techniques and captured complexity of the physical processes within the simulations will be quite challenging. Recent leading simulations are already extremely difficult to analyse due to their enormous size and complexity, and they will surely continue to grow. In fact, the next generation of supercomputers will grow more in the number of accessible CPUs than on the speedup of the individual CPUs and this fact will make the analysis of the simulations as challenging as performing the simulations itself. To keep a comparable level of accuracy, the interpretation of a simulation of the next generation of high precision experiments will need to massively involve virtual telescopes as described in the previous section. This will increase the need

of involving complex analysis pipelines for "observing" simulations, and might lead to a new branch of *virtual observers* in the astrophysics community, similar to the already, new formed branch of *computational astrophysicists*.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialised X-ray components in clusters of galaxies". Special thanks to Volker Springel for helping to improve the manuscript and providing Fig. 7 to Anna Watts for carefully reading the manuscript, to Martin Obergaulinger for very helpful discussions on Eulerian schemes, to Ewald Müller for the lecture notes on the Riemann problem and providing Fig. 10, to Elena Rasia for providing Fig. 16 and to Umberto Maio for providing Fig. 12. A.D. also gratefully acknowledges partial support from the PRIN2006 grant "Costituenti fondamentali dell'Universo" of the Italian Ministry of University and Scientific Research and from the INFN grant PD51.

References

T. Abel, P. Anninos, Y. Zhang, M.L. Norman, New Astron. **2**, 181 (1997)
O. Agertz, B. Moore, J. Stadel et al., Mon. Not. R. Astron. Soc. **380**, 963 (2007)
E. Athanassoula, A. Bosma, J.-C. Lambert, J. Makino, Mon. Not. R. Astron. Soc. **293**, 369 (1998)
J.S. Bagla, J. Astrophys. Astron. **23**, 185 (2002)
J.S. Bagla, S. Ray, New Astron. **8**, 665 (2003)
D.S. Balsara, J. Comput. Phys. **121**, 357 (1995)
J.M. Bardeen, J.R. Bond, N. Kaiser, A.S. Szalay, Astrophys. J. **304**, 15 (1986)
J. Barnes, P. Hut, Nature **324**, 446 (1986)
E. Bertschinger, COSMICS: Cosmological initial conditions and microwave anisotropy codes. astro-ph/9506070 (1995)
P. Bode, J.P. Ostriker, G. Xu, Astrophys. J. Suppl. Ser. **128**, 561 (2000)
A. Bonaldi, G. Tormen, K. Dolag, L. Moscardini, Mon. Not. R. Astron. Soc. **378**, 1248 (2007)
S. Borgani, K. Dolag, G. Murante et al., Mon. Not. R. Astron. Soc. **367**, 1641 (2006)
S. Borgani, D. Fabjan, L. Tornatore et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9322-7
G.L. Bryan, M.L. Norman, J.M. Stone, R. Cen, J.P. Ostriker, Comput. Phys. Commun. **89**, 149 (1995)
R. Cen, Astrophys. J. Suppl. Ser. **78**, 341 (1992)
S.C. Chapra, R.P. Canale, *Numerical Methods for Engineers: With Programming and Software Applications* (McGraw–Hill, New York, 1997)
P. Colella, H.M. Glaz, J. Comput. Phys. **59**, 264 (1985)
P. Colella, P.R. Woodward, J. Comput. Phys. **54**, 174 (1984)
H.M.P. Couchman, Astrophys. J. **368**, L23 (1991)
H.M.P. Couchman, P.A. Thomas, F.R. Pearce, Astrophys. J. **452**, 797 (1995)
R. Courant, K.O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience, New York, 1948)
W. Dehnen, Astrophys. J. **536**, L39 (2000)
K. Dolag, M. Jubelgas, V. Springel, S. Borgani, E. Rasia, Astrophys. J. Lett. **606**, L97 (2004)
K. Dolag, F. Vazza, G. Brunetti, G. Tormen, Mon. Not. R. Astron. Soc. **364**, 753 (2005)
K. Dolag, A.M. Bykov, A. Diaferio, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9319-2
G. Efstathiou, M. Davis, S.D.M. White, C.S. Frenk, Astrophys. J. Suppl. Ser. **57**, 241 (1985)
B. Einfeldt, SIAM J. Numer. Anal. **25**, 294 (1988)
B. Einfeldt, P.L. Roe, C.D. Munz, B. Sjogreen, J. Comput. Phys. **92**, 273 (1991)
S. Ettori, K. Dolag, S. Borgani, G. Murante, Mon. Not. R. Astron. Soc. **365**, 1021 (2006)
A.E. Evrard, Mon. Not. R. Astron. Soc. **235**, 911 (1988)
A.E. Evrard, T.J. MacFarland, H.M.P. Couchman et al., Astrophys. J. **573**, 7 (2002)
L.-L. Feng, C.-W. Shu, M. Zhang, Astrophys. J. **612**, 1 (2004)
C.S. Frenk, S.D.M. White, P. Bode et al., Astrophys. J. **525**, 554 (1999)
B. Fryxell, K. Olson, P. Ricker et al., Astrophys. J. Suppl. Ser. **131**, 273 (2000)
T. Fukushige, T. Ito, J. Makino et al., Publ. Astron. Soc. Jpn. **43**, 841 (1991)
D. Galli, F. Palla, Astron. Astrophys. **335**, 403 (1998)
A. Gardini, E. Rasia, P. Mazzotta et al., Mon. Not. R. Astron. Soc. **351**, 505 (2004)
R.A. Gingold, J.J. Monaghan, Mon. Not. R. Astron. Soc. **181**, 375 (1977)
F. Haardt, P. Madau, Astrophys. J. **461**, 20 (1996)
A. Harten, J. Comput. Phys. **49**, 357 (1983)
A. Harten, P.D. Lax, B. van Leer, SIAM Rev. **25**, 35 (1983)
A. Harten, B. Engquist, S. Osher, S.R. Chakravarthy, J. Comput. Phys. **71**, 231 (1987)

K. Heitmann, Z. Lukic, P. Fasel et al., astro-ph/0706.1270 (2007). Report LA-UR-07-1953
L. Hernquist, Astrophys. J. **404**, 717 (1993)
L. Hernquist, N. Katz, Astrophys. J. Suppl. Ser. **70**, 419 (1989)
R.W. Hockney, J.W. Eastwood, *Computer Simulation Using Particles* (Hilger, Bristol, 1988)
P. Hut, J. Makino, S. McMillan, Astrophys. J. **443**, L93 (1995)
I.T. Iliev, B. Ciardi, M.A. Alvarez et al., Mon. Not. R. Astron. Soc. **371**, 1057 (2006)
T. Ito, J. Makino, T. Fukushige et al., Publ. Astron. Soc. Jpn. **45**, 339 (1993)
G.-S. Jiang, C.-W. Shu, J. Comput. Phys. **126**, 202 (1996)
M. Jubelgas, V. Springel, K. Dolag, Mon. Not. R. Astron. Soc. **351**, 423 (2004)
J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310
H. Kang, J.P. Ostriker, R. Cen et al., Astrophys. J. **430**, 83 (1994)
W. Kapferer, T. Kronberger, J. Weratschnig, S. Schindler, Astron. Astrophys. **472**, 757 (2007)
N. Katz, D.H. Weinberg, L. Hernquist, Astrophys. J. Suppl. Ser. **105**, 19 (1996)
A. Kawai, T. Fukushige, J. Makino, M. Taiji, Publ. Astron. Soc. Jpn. **52**, 659 (2000)
A. Kravtsov, in *APS Meeting Abstracts* (2002). http://flux.aps.org/meetings/YR02/APR02/baps/abs/S820004.html
A.V. Kravtsov, A.A. Klypin, A.M. Khokhlov, Astrophys. J. Suppl. Ser. **111**, 73 (1997)
A.V. Kravtsov, D. Nagai, A.A. Vikhlinin, Astrophys. J. **625**, 588 (2005)
R. LeVeque, *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, 2002)
D. Levy, G. Puppo, G. Russo, Berkeley CPAM 762 (1999). arXiv:math/9911089
S. Li, J. Comput. Phys. **203**, 344 (2005)
S. Li, H. Li, R. Cen, Astrophys. J. Suppl. Ser. **174**, 1–12 (2008). astro-ph/0611863
L.B. Lucy, Astron. J. **82**, 1013 (1977)
U. Maio, K. Dolag, B. Ciardi, L. Tornatore, Mon. Not. R. Astron. Soc. **379**, 963 (2007)
J. Makino, Publ. Astron. Soc. Jpn. **43**, 621 (1991)
B.F. Mathiesen, A.E. Evrard, Astrophys. J. **546**, 100 (2001)
P. Mazzotta, E. Rasia, L. Moscardini, G. Tormen, Mon. Not. R. Astron. Soc. **354**, 10 (2004)
M. Meneghetti, P. Melchior, A. Grazian et al., Astron. Astrophys. (2008, in press). astro-ph/0711.3418
J.J. Monaghan, Annu. Rev. Astron. Astrophys. **30**, 543 (1992)
J.J. Monaghan, J. Comput. Phys. **136**, 298 (1997)
J.J. Monaghan, R.A. Gingold, J. Comput. Phys. **52**, 374 (1983)
J.P. Morris, J.J. Monaghan, J. Comput. Phys. **136**, 41 (1997)
J.F. Navarro, S.D.M. White, Mon. Not. R. Astron. Soc. **265**, 271 (1993)
B.W. O'Shea, K. Nagamine, V. Springel, L. Hernquist, M.L. Norman, Astrophys. J. Suppl. Ser. **160**, 1 (2005)
J.M. Owen, J.V. Villumsen, P.R. Shapiro, H. Martel, Astrophys. J. Suppl. Ser. **116**, 155 (1998)
P.J.E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, 1980)
U.-L. Pen, Astrophys. J. Suppl. Ser. **115**, 19 (1998)
K.G. Powell, P.L. Roe, T.J. Linde, T.I. Gombosi, D.L. de Zeeuw, J. Comput. Phys. **154**, 284 (1999)
C. Power, J.F. Navarro, A. Jenkins et al., Mon. Not. R. Astron. Soc. **338**, 14 (2003)
W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN. The Art of Scientific Computing*, 2nd edn. (Cambridge University Press, Cambridge, 1992)
T. Quinn, N. Katz, J. Stadel, G. Lake, Astrophys. J. (1997). Submitted but not published, astro-ph/9710043
E. Rasia, P. Mazzotta, S. Borgani et al., Astrophys. J. **618**, L1 (2005)
E. Rasia, S. Ettori, L. Moscardini et al., Mon. Not. R. Astron. Soc. **369**, 2013 (2006)
E. Rasia, P. Mazzotta, H. Bourdin et al., Astrophys. J. (2007, in press). astro-ph/0707.2614
P.M. Ricker, S. Dodelson, D.Q. Lamb, Astrophys. J. **536**, 122 (2000)
D. Ryu, J.P. Ostriker, H. Kang, R. Cen, Astrophys. J. **414**, 1 (1993)
H.-Y. Schive, C.-H. Chien, S.-K. Wong, Y.-C. Tsai, T. Chiueh, New Astron. (2007, in press). astro-ph/0707.2991
C.-W. Shu, in *Advanced Numerical Approximate of Nonlinear Hyperbolic Equations*, ed. by A. Quarteroni. Lecture Notes in Mathematics, vol. 1697 (Springer, Berlin, 1998), p. 285
D. Sijacki, V. Springel, T. di Matteo, L. Hernquist, Mon. Not. R. Astron. Soc. **380**, 877 (2007)
B.W. Silverman, *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability (Chapman & Hall, London, 1986)
V. Springel, Mon. Not. R. Astron. Soc. **364**, 1105 (2005)
V. Springel, L. Hernquist, Mon. Not. R. Astron. Soc. **333**, 649 (2002)
V. Springel, L. Hernquist, Mon. Not. R. Astron. Soc. **339**, 289 (2003)
V. Springel, M. White, L. Hernquist, Astrophys. J. **549**, 681 (2001a)
V. Springel, N. Yoshida, S.D.M. White, New Astron. **6**, 79 (2001b)

M. Steinmetz, Mon. Not. R. Astron. Soc. **278**, 1005 (1996a)
M. Steinmetz, in *IAU Symp. 171: New Light on Galaxy Evolution* (1996b), p. 259
F. Stoehr, S.D.M. White, V. Springel, G. Tormen, N. Yoshida, Mon. Not. R. Astron. Soc. **345**, 1313 (2003)
J.M. Stone, M.L. Norman, Astrophys. J. Suppl. Ser. **80**, 753 (1992)
R.S. Sutherland, M.A. Dopita, Astrophys. J. Suppl. Ser. **88**, 253 (1993)
G. Tormen, F.R. Bouchet, S.D.M. White, Mon. Not. R. Astron. Soc. **286**, 865 (1997)
S.D.M. White, in *Cosmology and Large Scale Structure*, ed. by R. Schaeffer, J. Silk, M. Spiro, J. Zinn-Justin (1996) p. 349
G. Xu, Astrophys. J. Suppl. Ser. **98**, 355 (1995)
Y.B. Zel'dovich, Astron. Astrophys. **5**, 84 (1970)

Chapter 13
Thermodynamical Properties of the ICM from Hydrodynamical Simulations

S. Borgani · A. Diaferio · K. Dolag · S. Schindler

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9317-4

Abstract Modern hydrodynamical simulations offer nowadays a powerful means to trace the evolution of the X-ray properties of the intra-cluster medium (ICM) during the cosmological history of the hierarchical build up of galaxy clusters. In this paper we review the current status of these simulations and how their predictions fare in reproducing the most recent X-ray observations of clusters. After briefly discussing the shortcomings of the self-similar model, based on assuming that gravity only drives the evolution of the ICM, we discuss how the processes of gas cooling and non-gravitational heating are expected to bring model predictions into better agreement with observational data. We then present results from the hydrodynamical simulations, performed by different groups, and how they compare with observational data. As terms of comparison, we use X-ray scaling relations between mass, luminosity, temperature and pressure, as well as the profiles of temperature and entropy. The results of this comparison can be summarised as follows: (*a*) simulations,

S. Borgani (✉)
Department of Astronomy, University of Trieste, via Tiepolo 11, 34143 Trieste, Italy
e-mail: borgani@oats.inaf.it

S. Borgani
INAF – National Institute for Astrophysics, Trieste, Italy

S. Borgani · A. Diaferio
INFN – National Institute for Nuclear Physics, Sezione di Trieste, Italy

A. Diaferio
e-mail: diaferio@ph.unito.it

A. Diaferio
Dipartimento di Fisica Generale "Amedeo Avogadro", Università degli Studi di Torino, Torino, Italy

K. Dolag
Max-Planck-Institut für Astrophysik, Karl-Schwarzschild Strasse 1, Garching bei München, Germany
e-mail: kdolag@mpa-garching.mpg.de

S. Schindler
Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstr. 25, 6020, Innsbruck, Austria
e-mail: sabine.schindler@uibk.ac.at

which include gas cooling, star formation and supernova feedback, are generally successful in reproducing the X-ray properties of the ICM outside the core regions; (*b*) simulations generally fail in reproducing the observed "cool core" structure, in that they have serious difficulties in regulating overcooling, thereby producing steep negative central temperature profiles. This discrepancy calls for the need of introducing other physical processes, such as energy feedback from active galactic nuclei, which should compensate the radiative losses of the gas with high density, low entropy and short cooling time, which is observed to reside in the innermost regions of galaxy clusters.

Keywords Cosmology: numerical simulations · Galaxies: clusters · Hydrodynamics · X-ray: galaxies

1 Introduction

Clusters of galaxies form from the collapse of exceptionally high density perturbations having typical size of ~ 10 Mpc in a comoving frame. As such, they mark the transition between two distinct regimes in the study of the formation of cosmic structures. The evolution of structures involving larger scales is mainly driven by the action of gravitational instability of the dark matter (DM) density perturbations and, as such, it retains the memory of the initial conditions. On the other hand, galaxy-sized structures, which form from initial fluctuations on scales of ~ 1 Mpc, evolve under the combined action of gravity and of complex gas-dynamical and astrophysical processes. On such scales, gas cooling, star formation and the subsequent release of energy and metal feedback from supernovae (SN) and active galactic nuclei (AGN) have a deep impact on the observational properties of the diffuse gas and of the galaxy population.

In this sense, clusters of galaxies can be used as invaluable cosmological tools and astrophysical laboratories (see Rosati et al. 2002 and Voit 2005 for reviews). These two aspects are clearly interconnected with each other. From the one hand, the evolution of the population of galaxy clusters and their overall baryonic content provide in principle powerful constraints on cosmological parameters. On the other hand, for such constraints to be robust, one has to understand in detail the physical properties of the intra-cluster medium (ICM) and its interaction with the galaxy population.

The simplest model to predict the properties of the ICM and their evolution has been proposed by Kaiser (1986). This model is based on the assumption that the evolution of the thermodynamical properties of the ICM is determined only by gravity, with gas heated to the virial temperature of the hosting DM halos by accretion shocks (Bykov et al. 2008—Chap. 7, this issue). Since gravitational interaction does not introduce any preferred scale, this model has been called "self-similar"[1]. As we shall discuss, this model provides precise predictions on the shape and evolution of scaling relations between X-ray luminosity, entropy, total and gas mass, which have been tested against numerical hydrodynamical simulations (e.g., Eke et al. 1998, Bryan and Norman 1998). These predictions have been recognised for several years to be at variance with a number of observations. In particular, the observed relation between X-ray luminosity and temperature (e.g., Markevitch 1998; Arnaud and Evrard 1999;

[1]Strictly speaking, self-similarity also requires that no characteristic scales are present in the underlying cosmological model. This means that the Universe must obey the Einstein–de-Sitter expansion law and that the shape of the power spectrum of density perturbations is a featureless power law. In any case, the violation of self-similarity introduced by the standard cosmological model is negligible with respect to that related to the non-gravitational effects acting on the gas.

Osmond and Ponman 2004) is steeper and the measured level of gas entropy higher than expected (e.g., Ponman et al. 2003; Pratt and Arnaud 2005), especially for poor clusters and groups. This led to the concept that more complex physical processes, related to the heating from astrophysical sources of energy feedback, and radiative cooling of the gas in the central cluster regions, play a key role in determining the properties of the diffuse hot baryons.

Although semi-analytical approaches (e.g., Tozzi and Norman 2001; Voit 2005, and references therein) offer invaluable guidelines to this study, it is only with hydrodynamical simulations that one can capture the full complexity of the problem, so as to study in detail the existing interplay between cosmological evolution and the astrophysical processes.

In the last years, ever improving code efficiency and supercomputing capabilities have opened the possibility to perform simulations over fairly large dynamical ranges, thus allowing to resolve scales of a few kiloparsecs (kpc), which are relevant for the formation of single galaxies, while capturing the global cosmological environment on scales of tens or hundreds of Megaparsecs (Mpc), which are relevant for the evolution of galaxy clusters. Starting from first attempts, in which only simplified heating schemes were studied (e.g., Navarro et al. 1995; Bialek et al. 2001; Borgani et al. 2002), a number of groups have studied the effect of introducing also cooling (e.g., Katz and White 1993; Lewis et al. 2000; Muanwong et al. 2001; Davé et al. 2002; Tornatore et al. 2003), of more realistic sources of energy feedback (e.g., Borgani et al. 2004; Kay et al. 2007; Nagai et al. 2007a; Sijacki et al. 2007), of thermal conduction (e.g., Dolag et al. 2004), and of non-thermal pressure support from magnetic fields (e.g., Dolag et al. 2001) and cosmic rays (e.g., Pfrommer et al. 2007).

In this paper, we will review the recent advancement performed in this field of computational cosmology and critically discuss the comparison between simulation predictions and observations, by restricting the discussion to the thermal effects. As such, this paper complements the reviews by Borgani et al. (2008—Chap. 18, this issue), which reviews the study of the ICM chemical enrichment and by Dolag et al. (2008b—Chap. 15, this issue), which reviews the study of the non-thermal properties of the ICM from simulations. We refer to the reviews by Dolag et al. (2008a—Chap. 12, this issue) for a description of the techniques of numerical simulations and by Kaastra et al. (2008—Chap. 9, this issue) for an overview of the observed thermal properties of the ICM.

The scheme of the presentation is as follows. In Sect. 2 we briefly discuss the self-similar model of the ICM and how the action of non-gravitational heating and cooling are expected to alter the predictions of this model. Sections 3 and 4 overview the results obtained on the comparison between observed and simulated scaling relations and profiles of X-ray observable quantities, respectively. In Sect. 5 we summarise and critically discuss the results presented.

2 Modelling the ICM

2.1 The Self-Similar Scaling

The simplest model to predict the observable properties of the ICM is based on the assumption that gravity only determines the thermodynamical properties of the hot diffuse gas (Kaiser 1986). Since gravity does not have preferred scales, we expect clusters of different sizes to be the scaled version of each other. This is the reason why this model has been called self-similar.

If, at redshift z, we define M_{Δ_c} to be the mass contained within the radius r_{Δ_c}, encompassing a mean density Δ_c times the critical density $\rho_c(z)$, then $M_{\Delta_c} \propto \rho_c(z)\Delta_c r^3_{\Delta_c}$. The

critical density of the universe scales with redshift as $\rho_{\rm c}(z) = \rho_{\rm c0} E^2(z)$, where $E(z)$ is given by

$$E(z) = \left[(1+z)^3 \Omega_{\rm m} + (1+z)^2 \Omega_{\rm k} + \Omega_\Lambda\right]^{1/2}, \tag{1}$$

where $\Omega_{\rm m}$ and Ω_Λ are the density parameters associated to the non-relativistic matter and to the cosmological constant, respectively, $\Omega_{\rm k} = 1 - \Omega_{\rm m} - \Omega_\Lambda$ and we neglect any contribution from relativistic species.

Therefore, the cluster size $r_{\Delta_{\rm c}}$ scales with z and $M_{\Delta_{\rm c}}$ as $r_{\Delta_{\rm c}} \propto M_{\Delta_{\rm c}}^{1/3} E^{-2/3}(z)$, so that, assuming hydrostatic equilibrium, cluster mass scales with temperature T as

$$M_{\Delta_{\rm c}} \propto T^{3/2} E^{-1}(z). \tag{2}$$

If $\rho_{\rm gas}$ is the gas density, the corresponding X-ray luminosity is

$$L_X = \int_V \left(\frac{\rho_{\rm gas}}{\mu m_{\rm p}}\right)^2 \Lambda(T)\, {\rm d}V, \tag{3}$$

where $\Lambda(T) \propto T^{1/2}$ for pure thermal Bremsstrahlung emission. If gas accretes along with DM by gravitational instability during the formation of the cluster halo, then we expect that $\rho_{\rm gas}(r) \propto \rho_{\rm M}(r)$, so that

$$L_X \propto M_{\Delta_{\rm c}} \rho_{\rm c} T^{1/2} \propto T^2 E(z). \tag{4}$$

Another useful quantity characterising the thermodynamical properties of the ICM is the entropy (Voit 2005) which, in X-ray studies of the ICM, is usually defined as

$$K = \frac{{\rm k_B} T}{\mu m_{\rm p} \rho_{\rm gas}^{2/3}}, \tag{5}$$

where $\rm k_B$ is the Boltzmann constant, μ the mean molecular weight ($\simeq 0.58$ for a plasma of primordial composition) and $m_{\rm p}$ the proton mass. With the above definition, the quantity K is the constant of proportionality in the equation of state of an adiabatic mono-atomic gas, $P = K \rho_{\rm gas}^{5/3}$. Using the thermodynamic definition of specific entropy, $s = c_V \ln(P/\rho_{\rm gas}^{5/3})$ (c_V: heat capacity at constant volume), one obtains $s = {\rm k_B} \ln K^{3/2} + s_0$, where s_0 is a constant. Another quantity, often called "entropy" in the cluster literature, which we will also use in the following, is

$$S = {\rm k_B} T n_{\rm e}^{-2/3}, \tag{6}$$

where $n_{\rm e}$ is the electron number density. According to the self-similar model, this quantity, computed at a fixed overdensity $\Delta_{\rm c}$, scales with temperature and redshift according to

$$S_{\Delta_{\rm c}} \propto T (1+z)^{-2}. \tag{7}$$

As already mentioned in the introduction, a number of observational facts from X-ray data point against the simple self-similar picture. The steeper slope of the L_X–T relation (Markevitch 1998; Arnaud and Evrard 1999; Osmond and Ponman 2004), $L_X \propto T^\alpha$ with $\alpha \simeq 3$ for clusters and possibly larger for groups, the excess entropy in poor clusters and groups (Ponman et al. 2003; Pratt and Arnaud 2005; Piffaretti et al. 2005) and the decreasing trend of the gas mass fraction in poorer systems (Lin et al. 2003; Sanderson et al. 2003) all point toward the presence of some mechanism which significantly affects the ICM thermodynamics.

2.2 Heating and Cooling the ICM

The first mechanism, that has been introduced to break the ICM self-similarity, is non-gravitational heating (e.g., Evrard and Henry 1991; Kaiser 1991; Tozzi and Norman 2001). The idea is that by increasing the gas entropy with a given extra heating energy per gas particle $E_{\rm h}$ prevents gas from sinking to the centre of DM halos, thereby reducing gas density and X-ray emissivity. This effect will be large for small systems, whose virial temperature is $k_B T \lesssim E_{\rm h}$, while leaving rich clusters with $k_B T \gg E_{\rm h}$ almost unaffected. Therefore, we expect that the X-ray luminosity and gas content are relatively more suppressed in poorer systems, thus leading to a steepening of the L_X–T relation.

The notion of non-gravitational heating has been first implemented in non–radiative (i.e. neglecting the effect of cooling) hydrodynamical simulations by either injecting entropy in an impulsive way at a given redshift (Navarro et al. 1995; Bialek et al. 2001), or by adding energy in a redshift-modulated way, so as to mimic the rate of SN explosions from an external model of galaxy formation (Borgani et al. 2002). In Fig. 1 we show the different efficiency that different heating mechanisms have in breaking the self-similar behaviour of the entropy profiles in objects of different mass, ranging from a Virgo-like cluster to a poor galaxy group. According to the self-similar model, the profiles of reduced entropy, S/T, should be independent of the cluster mass. This is confirmed by the left panel, which also shows that these profiles have a slope consistent with that predicted by a model in which gas is shock heated by spherical accretion in a DM halo, under the effect of gravity only (Tozzi and Norman 2001). The central panel shows instead the effect of adding energy from SN, whose rate is that predicted by a semi-analytical model of galaxy formation. In this case, which corresponds to a total heating energy of about 0.3 keV/particle, the effect of extra heating starts being visible, but only for the smaller system. It is only with the pre–heating scheme, based on imposing an entropy floor of 50 keV cm^2, that self-similarity is clearly broken. While this heating scheme is effective in reproducing the observed L_X–T relation, it produces large isentropic cores, a prediction which is at variance with respect to observations (e.g., Donahue et al. 2006).

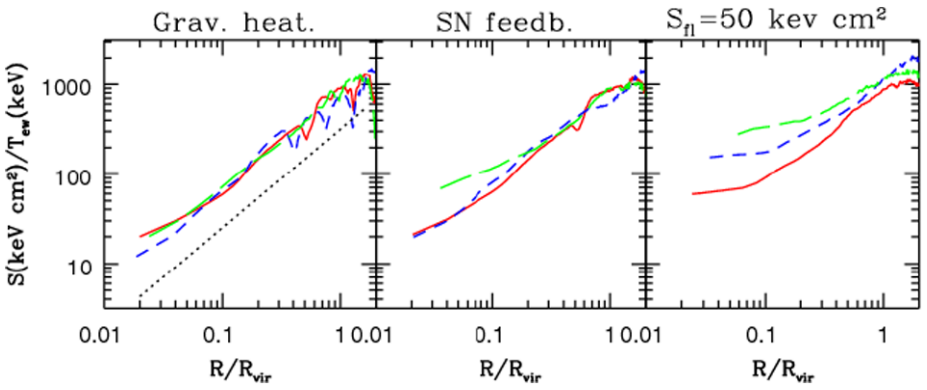


Fig. 1 Profiles of reduced entropy, S/T, for non-radiative simulations (from Borgani et al. 2001). The *left panel* is for simulations including only gravitational heating, the *central panel* is for runs including SN feedback as predicted by a semi-analytical model of galaxy formation and the *right panel* is with pre-heating with an entropy threshold at redshift $z = 3$. Solid, *short-dashed* and *long-dashed* curves are for a 3 keV cluster, for a 1 keV group and for a 0.5 keV group, respectively. The *dotted straight line* in the *left panel* shows the analytical prediction by Tozzi and Norman (2001) for the entropy profile associated to gravitational heating

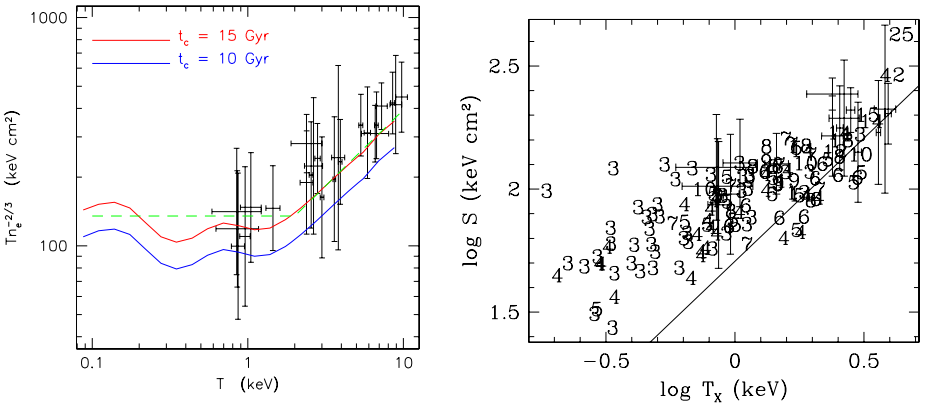


Fig. 2 *Left panel*: the relation between entropy and temperature for gas having a fixed value of the cooling time (from Voit and Bryan 2001). The *crosses with error bars* are observational data on the entropy measured at one-tenth of the virial radius for clusters and groups by Ponman et al. (1999). *Right panel*: a comparison between observations (*error bars with crosses*; Ponman et al. (1999) and simulations, including radiative cooling and star formation (*numbers*), for the entropy in the central regions of galaxy groups and clusters. The *solid line* shows the prediction of the self-similar model (from Davé et al. 2002)

Although it may look like a paradox, radiative cooling has been also suggested as a possible alternative to non-gravitational heating to increase the entropy level of the ICM and suppressing the gas content in poor systems. As originally suggested by Voit and Bryan (2001), cooling provides a selective removal of low-entropy gas from the hot X-ray emitting phase (see also Wu and Xue 2002). As a consequence, while the global entropy of the baryons decreases, the entropy of the X-ray emitting gas increases. This is illustrated in the left panel of Fig. 2 (from Voit and Bryan 2001). In this plot, each of the two curves separates the upper portion of the entropy–temperature plane, where the gas has cooling time larger than the age of the system, from the lower portion, where gas with short cooling time resides. This implies that only gas having a relatively high entropy will be observed as X-ray emitting, while the low-entropy gas will be selectively removed by radiative cooling. The comparison with observational data of clusters and groups, also reported in this plot, suggests that their entropy level may well be the result of this removal of low-entropy gas operated by radiative cooling. This analytical prediction has been indeed confirmed by radiative hydrodynamical simulations. The right panel of Fig. 2 shows the results of the simulations by Davé et al. (2002) on the temperature dependence of the central entropy of clusters and groups. Quite apparently, the entropy level in simulations is well above the prediction of the self-similar model, by a relative amount which increases with decreasing temperature, and in reasonable agreement with the observed entropy level of poor clusters and groups.

Although cooling may look like an attractive solution, it suffers from the drawback that a too large fraction of gas is converted into stars in the absence of a source of heating energy which regulates the cooling runaway. Indeed, while observations indicate that only about 10 per cent of the baryon content of a cluster is in the stellar phase (e.g., Balogh et al. 2001; Lin et al. 2003), radiative simulations, like those shown in Fig. 2, convert into stars up to ~ 50 per cent of the gas.

Another paradoxical consequence of cooling is that it increases the temperature of the hot X-ray emitting gas at the centre of clusters. This is shown in the left panel of Fig. 3 (from Tornatore et al. 2003), which compares the temperature profiles for the non-radiative run of a Virgo-like cluster with a variety of radiative runs, based on different ways of supplying

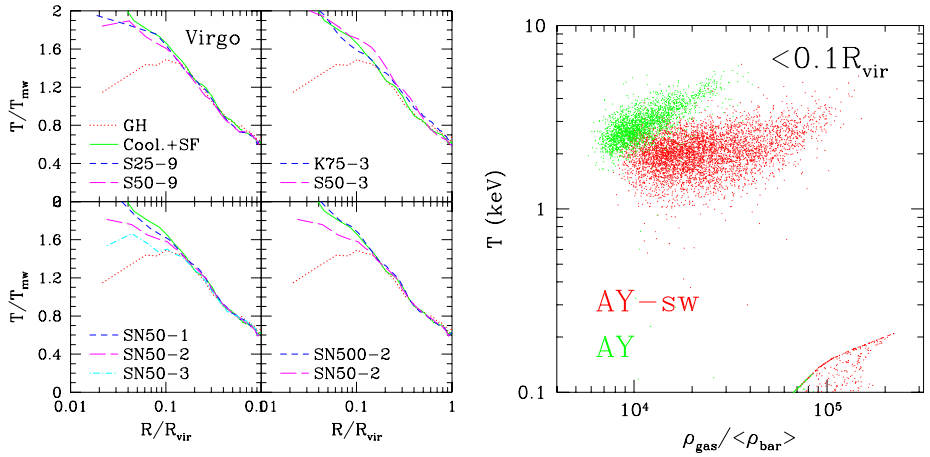


Fig. 3 *Left panel*: temperature profiles from hydrodynamical simulations of a ~ 3 keV galaxy cluster. In all panels the *dotted* and the *solid curves* correspond to a non-radiative run and to a run including cooling and star formation. The other curves are for different recipes of gas heating (from Tornatore et al. 2003). *Right panel*: the relation between temperature and overdensity for gas particles within $0.1r_{200}$ for SPH simulations of a cluster of mass $\simeq 10^{14}\, h^{-1}\mathrm{M}_{\odot}$. *Upper* (*green*) *points* are for a run which includes feedback through galactic winds with a velocity of 500 km s^{-1}, while the *lower* (*red*) *points* are for a run based on assuming stronger winds (sw), with a twice as large velocity. Both runs include a model of chemical enrichment (Borgani et al. 2008—Chap. 18, this issue) which assumes an Initial Mass Function for star formation by Arimoto and Yoshii 1987 (AY). The *points in the bottom right corner* are star-forming gas particles

non-gravitational heating. The effect of introducing cooling is clearly that of steepening the temperature profiles in the core regions, while leaving it unchanged at larger radii. The reason for this is that cooling causes a lack of central pressure support. As a consequence, gas starts flowing in sub-sonically from more external regions, thereby being heated by adiabatic compression. As we shall discuss in Sect. 4, this feature of cooling makes it quite difficult to reproduce the structure of the cool cores observed in galaxy clusters.

Steepening of the central temperature profiles and overcooling are two aspects of the same problem. In principle, the solution to this problem should be provided by a suitable scheme of gas heating which regulates star formation, while maintaining pressurised gas in the hot phase. The right panel of Fig. 3 compares the temperature–density phase diagrams for gas particles lying in the central region of an SPH-simulated cluster, when using two different feedback efficiencies. The two simulations include cooling, star formation and feedback in the form of galactic winds powered by SN explosions, following the scheme introduced by Springel and Hernquist (2003a). The upper (green) and the lower (red) clouds of high temperature particles correspond to a wind velocity of 500 km s^{-1} and of 1000 km s^{-1}, respectively. This plot illustrates another paradoxical effect: in the same way that cooling causes an increase of the temperature of the hot phase, supplying energy with an efficient feedback causes a decrease of the temperature. The reason for this is that extra energy compensates radiative losses, thereby maintaining the pressure support for gas which would otherwise have a very short cooling time, thereby allowing it to survive on a lower adiabat. It is also worth reminding that cooling efficiency increases with the numerical resolution (e.g., Balogh et al. 2001; Borgani et al. 2006). Therefore, for a feedback mechanism to work properly, it should be able to stabilise the cooling efficiency in a way which is independent of resolution.

In the light of these results, it is clear that the observed lack of self-similarity in the X-ray properties of clusters cannot be simply explained on the grounds of a single effect. The emerging picture is that the action of cooling and of feedback energy, e.g. associated to SN explosions and AGN, should combine in a self-regulated way. As we shall discuss in the following sections, hydrodynamical simulations of galaxy clusters in a cosmological context demonstrate that achieving this heating/cooling balance is not easy and represents nowadays one of the most challenging tasks in the numerical study of clusters.

As an example, we show in Fig. 4 how the gas density of a simulated cluster changes, both at $z = 2$ (left panels) and at $z = 0$ (right panels), when cooling and star formation are combined with different forms of non-gravitational heating (from Borgani et al. 2005). The comparison of the top and central panels shows the effect of increasing the kinetic energy carried by galactic outflows by a factor of six. The stronger winds are quite efficient in stopping star formation in the small halos, which are washed out, and make the larger ones slightly puffier, while preserving the general structure of the cosmic web surrounding the Lagrangian cluster region. Comparing the top and the bottom panels shows instead the effect of adding to galactic winds also the effect of an entropy floor. Although the energy budget of the feedback schemes of the central and bottom panels are quite comparable, the effect on the gas distribution is radically different. Imposing an entropy floor at $z = 3$ with an impulsive heating generates a much smoother gas density distribution, both at $z = 2$ and at $z = 0$. In this case, the filamentary structure of the gas distribution is completely erased, while only the largest halos are able to retain part of their gas content. This demonstrates that a fixed amount of energy feedback can provide largely different results on the ICM thermodynamical properties, depending on the epoch and on the density at which it is released.

3 Scaling Relations

So far, we have qualitatively discussed how simple models of pre-heating and radiative cooling can reproduce the observed violation of self-similarity in the X-ray properties of galaxy clusters. In this and in the following sections we will focus the discussion on a more detailed comparison between simulation results and observational data, and on the implications of this comparison on our current understanding of the feedback mechanisms which regulate star formation and the evolution of the galaxy population. As a starting point for the comparison between observed and simulated X-ray cluster properties, we describe how observable quantities are computed from hydrodynamical simulations and how they compare to the analogous quantities derived from observational data.

As for the X-ray luminosity, it is computed by summing the contributions to the emissivity, ε_i, carried by all the gas elements (particles in a SPH run and cells in an Eulerian grid-based run), $L_X = \sum_i \varepsilon_i$, where the sum extends over all the gas elements within the region where L_X is computed. The contribution from the i-th gas element is usually written as

$$\varepsilon_i = n_{\mathrm{e},i} n_{\mathrm{H},i} \Lambda(T_i, Z_i)\, \mathrm{d}V_i, \tag{8}$$

where $n_{\mathrm{e},i}$ and $n_{\mathrm{H},i}$ are the number densities of electrons and of hydrogen atoms, respectively, associated to the i-th gas element of given density ρ_i, temperature T_i and metallicity Z_i. Furthermore, $\Lambda(T, Z)$ is the temperature- and metallicity-dependent cooling function (e.g., Sutherland and Dopita 1993) computed within a given energy band, while $\mathrm{d}V_i = m_i/\rho_i$ is the volume of the i-th gas element, having mass m_i.

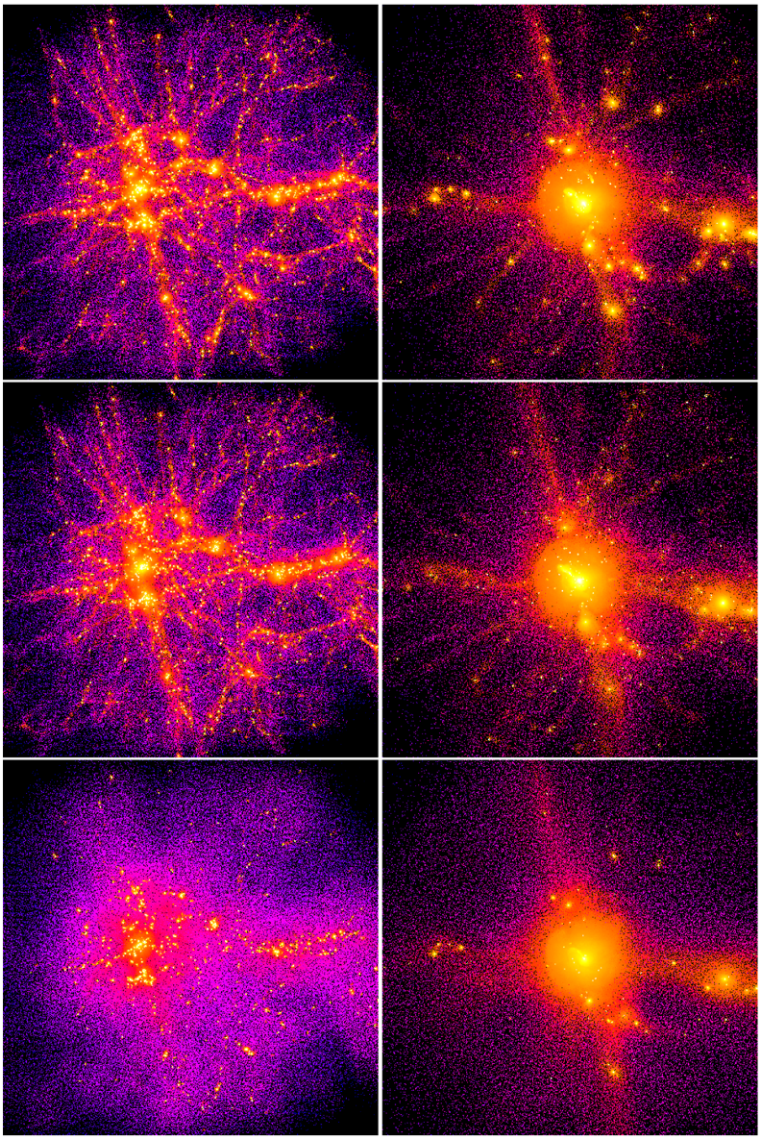


Fig. 4 The maps of the gas density for simulations of a Virgo-like cluster at $z = 2$ and $z = 0$ (*left* and *right panels*, respectively), including cooling, star formation and different forms of non-gravitational heating. *Upper panels* are for a run which includes galactic winds with a velocity of about 340 $\mathrm{km\,s^{-1}}$, the *central panels* is for galactic winds with velocity of about 830 $\mathrm{km\,s^{-1}}$ and the *bottom panels* is for galactic winds as in the *top panels*, but also adding a pre-heating with an entropy threshold of 100 $\mathrm{keV\,cm^2}$ at $z = 3$. At $z = 0$ the size of the box is of $11.7h^{-1}$ Mpc, while at $z = 2$ is corresponds to $17.5h^{-1}$ Mpc comoving (from Borgani et al. 2005)

As for the temperature, different proxies to its X-ray observational definition have been proposed in the literature, which differ from each other in the expression for the weight

assigned to each gas element. In general, the ICM temperature can be written as

$$T = \frac{\sum_i w_i T_i}{\sum_i w_i}, \tag{9}$$

where T_i is the temperature of the i-gas element, which contributes with the weight w_i. The mass-weighted definition of temperature, $T_{\rm mw}$, is recovered for $w_i = m_i$ (m_i: mass of the i-th gas element), which also coincides with the electron temperature $T_{\rm e}$ for a fully ionised plasma. A more observation-oriented estimate of the ICM temperature is provided by the emission-weighted definition, $T_{\rm ew}$, which is obtained for $w_i = \varepsilon_i$ (e.g., Evrard et al. 1996). The idea underlying this definition is that each gas element should contribute to the overall spectrum according to its emissivity.

Mazzotta et al. (2004) pointed out that the thermal complexity of the ICM is such that the overall spectrum is given by the superposition of several single-temperature spectra, each one associated to one thermal phase. In principle, the superposition of several single-temperature spectra cannot be described by a single-temperature spectrum. However, when fitting it to a single-temperature model in a typical finite energy band, where X-ray telescopes are sensitive, the cooler gas phases are relatively more important in providing the high-energy cut-off of the spectrum and, therefore, in determining the temperature resulting from the spectral fit. In order to account for this effect, Mazzotta et al. (2004) introduced a spectroscopic-like temperature, $T_{\rm sl}$, which is recovered from (9) by using the weight $w_i = \rho_i m_i T^{\alpha-3/2}$. By using $\alpha = 0.75$, this expression for $T_{\rm sl}$ was shown to reproduce within few percent the temperature obtained from the spectroscopic fit, at least for clusters with temperature above 2–3 keV. More complex fitting expressions have been provided by Vikhlinin (2006), who generalised the spectroscopic-like temperature to the cases of lower temperature and arbitrary metallicity.

3.1 The Luminosity–Temperature Relation

The L_X–T relation represented the first observational evidence against the self-similar model. This relation has been shown by several independent analyses to have a slope, $L_X \propto T^\alpha$, with $\alpha \simeq 3$ for $T \gtrsim 2$ keV (e.g., White et al. 1997), with indications for a flattening to $\alpha \lesssim 2.5$ for the most massive systems (Allen and Fabian 1998). The scatter in this relation is largely contributed by the cool-core emission, so that it significantly decreases when excising the cores (Markevitch 1998) or removing cool-core systems (Arnaud and Evrard 1999). A change of behaviour is also observed at the scales of groups, $T \lesssim 2$ keV, which generally displays a very large scatter (Osmond and Ponman 2004).

Hydrodynamical simulations by Bialek et al. (2001) and by Borgani et al. (2002) demonstrated that simple pre-heating models, based on the injection of entropy at relatively high redshift, can reproduce the observed slope of the L_X–T relation. Davé et al. (2002) showed that a similar result can also be achieved with simulations including cooling only, the price to be paid being a large overcooling. Muanwong et al. (2002) and Tornatore et al. (2003) demonstrated that combining cooling with pre-heating models can eventually decrease the total amount of stars to an acceptable level, while still providing a slope of the L_X–T relation close to the observed one.

In Fig. 5 we show the comparison between observations and simulations in which the non-gravitational gas heating assumes the energy budget made available by the SN explosions, whose rate is computed from the simulated star formation rate. The left panel shows the results from the SPH GADGET simulations by Borgani et al. (2004). These simulations

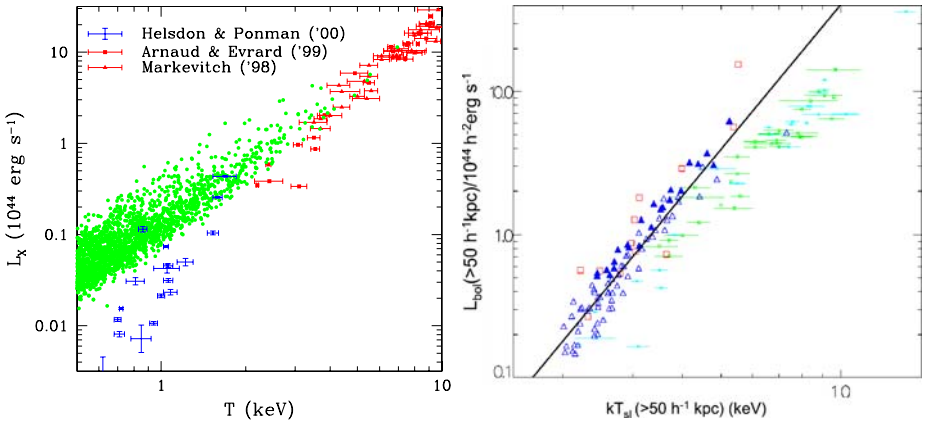


Fig. 5 *Left panel*: The relation between bolometric X-ray luminosity and emission-weighted temperature for simulations by Borgani et al. (2004), including cooling, star formation and feedback in the form of galactic winds (*green circles*), compared with observational data for clusters (Arnaud and Evrard 1999; Markevitch 1998) and groups (Helsdon and Ponman 2000). *Right panel*: the relation between X-ray luminosity, estimated outside the core regions, and the spectroscopic-like temperature, for simulations which include cooling, star formation and a form of "targeted" feedback (from Kay et al. 2007). Observational data for clusters are the same as in the *left panel*

included the same model of kinetic feedback used in the simulations shown in the top panel of Fig. 4. This feedback model was shown by Springel & Hernquist (2003b) to be quite successful in producing a cosmic star formation history similar to the observed one. Quite apparently, these simulations provided a reasonable relation at the scale of clusters, $T \gtrsim 3$ keV, while failing to produce slope and scatter at the scale of groups.

The right panel show the results by Kay et al. (2007), which these authors plotted only for systems with $T > 2$ keV. Kay et al. also used SPH simulations based on the GADGET code, but with a different feedback scheme. In this scheme, energy made available by SN explosions is assigned in a "targeted" way to suitably chosen gas particles, which surround the star-forming regions, so that their entropy is raised in such a way to prevent them from cooling. Therefore, while the amount of energy is self-consistently computed from star formation, the way in which it is thermalised to the gas is suitably tuned. These simulations predict a too high normalisation of the L_X–T relation. This result is interpreted by Kay et al. (2007) as due to the fact that their simulations produce too low temperatures for clusters of a given mass, as a consequence of the incorrect cool-core structure.

Besides the slope of the local L_X–T relation, also its evolution carries information about the thermodynamical history of the ICM. Thanks to the increasing statistics of distant clusters observed in the last years with Chandra and XMM-Newton, a number of authors analysed this evolution out to the highest redshifts, $z \simeq 1.3$, where clusters have been detected so far. These analyses generally indicate that the amplitude of the L_X–T relation has a positive evolution out to $z \simeq 0.5$–0.6 (e.g., Arnaud et al. 2002; Lumb et al. 2004; Kotov and Vikhlinin 2005), with hints for a possible inversion of this trend at higher redshift (e.g., Ettori et al. 2004b; Maughan et al. 2006; Branchesi et al. 2007). In the attempt of interpreting these results, Voit (2005) showed that radiative cooling, combined with a modest amount of pre-heating with extra entropy, predicts an evolution of the L_X–T relation slower than that of the self-similar model, also with an inversion of the trend at high redshift. Although in qualitative agreement with observations, these models have however difficulties in following the observed positive evolution at $z \lesssim 0.5$ (e.g., see Fig. 14 by Maughan et al. 2006).

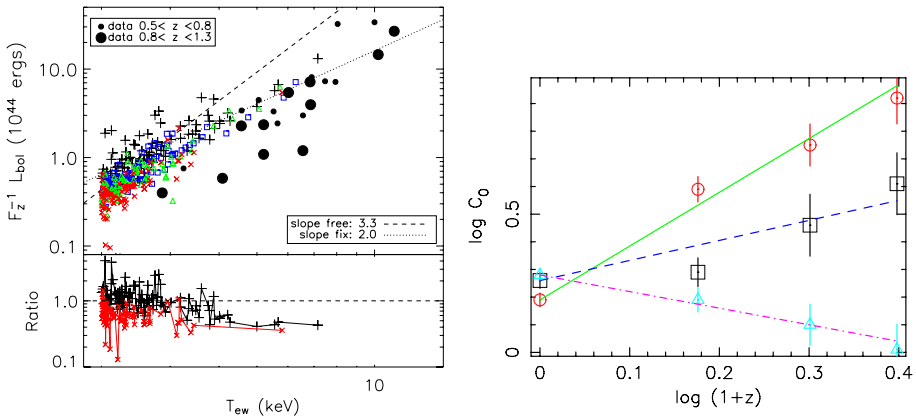


Fig. 6 *Left panel*: A comparison of the evolution of the L_X–T relation for the simulated clusters by Borgani et al. (2004) and the Chandra data analysed by Ettori et al. (2004b) (from Ettori et al. 2004a). The "*plus*" *symbols* are the relation at $z = 0$, with the *dotted* and the *dashed lines* providing the best-fit relations when the slope is kept fixed at 2 and is left free, respectively. The *squares*, the *triangles* and the *crosses* show the simulation results at $z = 0.5, 0.7$ and 1, respectively. The *small* and the *big dots* show the observational data at $0.5 < z < 0.8$ and at $z > 0.8$, respectively. The *lower panel* shows the ratio between the measured luminosities at $z = 0$ ("*plus*" *symbols*) and at $z = 1$ (*crosses*) and the corresponding best-fitting power laws. *Right panel*: normalisation of the L_X–T relation as a function of redshift, for the simulated clusters by Muanwong et al. (2006) in the radiative run with no feedback (*solid line*), in a run including an impulsive pre-heating at $z = 4.5$ (*dashed line*), and in a run including a form of "targeted" feedback (*dot-dashed line*)

As for the comparison with numerical predictions, Ettori et al. (2004a) analysed the evolution of the L_X–T relation from the radiative simulation by Borgani et al. (2004), which includes the effect of SN feedback in the form of galactic winds. As a result, they found that the normalisation of this relation for the simulated clusters at high redshift is higher than for real clusters (see left panel of Fig. 6). Muanwong et al. (2006) analysed three sets of simulated clusters, based on radiative cooling only, on gas pre-heating at high redshift and on the same "targeted" SN feedback model used by Kay et al. (2007). As shown in the right panel of Fig. 6, they found that these three models predicts rather different evolutions of the L_X–T relation, thus confirming it to be a sensitive test for the thermodynamical history of the ICM. However, so far none of the numerical models proposed is able to account for the observational indication for an inversion of the L_X–T evolution at $z \simeq 0.5$. This represents still an open issue, whose implications for a realistic modelling of the ICM physics are still to be understood.

3.2 The Mass–Temperature Relation

The relation between total collapsed mass and temperature has received much consideration both from the observational and the theoretical side, in view of its application for the use of galaxy clusters as tools to measure cosmological parameters (e.g., Voit 2005; Borgani 2007). The relation between ICM temperature and total mass should be primarily dictated by the condition of hydrostatic equilibrium. For this reason, the expectation is that this relation should have a rather small scatter and be insensitive to the details of the heating/cooling processes. For a spherically symmetric system, the condition of hydrostatic

equilibrium translates into the mass estimator (e.g. Sarazin 1988)

$$M(<r) = -\frac{r k_B T(r)}{G\mu m_p}\left[\frac{d \ln \rho_{gas}(r)}{d \ln r} + \frac{d \ln T(r)}{d \ln r}\right]. \quad (10)$$

Here $M(<r)$ is the total mass within the cluster-centric distance r, while $T(r)$ is the temperature measured at r. As shown in the previous section, the processes of heating/cooling are expected to modify the gas thermodynamics only in the central cluster regions, while the bulk of the ICM is dominated by gravitational processes. This implies that, in principle, the total mass estimate from (10) should be rather stable. However, since the X-ray spectroscopic temperature is sensitive to the thermal complexity of the ICM in the central regions (e.g., Mazzotta et al. 2004), a change of temperature and density profile in these regions may translate into a sizable effect in the mass–temperature relation provided by (10).

Equation (10) has been often applied in the literature by modelling the gas density profile with a β-model (Cavaliere and Fusco-Femiano 1976),

$$\rho_{gas}(r) = \frac{\rho_0}{[1 + (r/r_c)^2]^{3\beta/2}}, \quad (11)$$

where r_c is the core radius, and assuming either isothermal gas or a polytropic gas, $\rho_{gas} \propto T^{\gamma-1}$, to account for the presence of temperature gradients. In this case, (10) can be recast in the form

$$M(<r) \simeq 1.1 \times 10^{14} \beta\gamma \frac{T(r)}{\mathrm{keV}} \frac{r}{h^{-1}\,\mathrm{Mpc}} \frac{(r/r_c)^2}{1 + (r/r_c)^2}. \quad (12)$$

Based on the analysis of non-radiative cluster simulations, Schindler (1996) and Evrard et al. (1996) argued that the X-ray temperature provides a rather precise determination of the cluster mass, with an intrinsic scatter of only about 15 per cent at r_{500}. Finoguenov et al. (2001) applied (12) to ROSAT imaging and ASCA spectroscopic data (see also Nevalainen et al. 2000). They found that the resulting M–T relation has a normalisation about 40 per cent lower than that from the simulations by Evrard et al. (1996). Although introducing the effect of cooling, star formation and SN feedback provides a 20 per cent lower normalisation, this was not yet enough to recover agreement with observations. Independent analyses (Muanwong et al. 2002; Borgani et al. 2004) showed that applying to simulated clusters (12), which is used to estimate masses of real clusters, leads to a mass underestimate of about 20 per cent. This bias in the mass estimate is enough to bring the simulated and the observed M–T relations into reasonable agreement. These analyses were still based on the emission-weighted temperature in the simulation analysis.

Rasia et al. (2005) showed that using the spectroscopic-like definition, T_{sl}, leads to a mass underestimate of up to ~ 30 per cent with respect to the true cluster mass. This result is shown in Fig. 7. The left panel reports the relation between T_{sl} for simulated clusters and the true total cluster mass, also compared with the observed M–T relation by Finoguenov et al. (2001). Quite apparently, the relation from simulations lies well above the observational one. However, this difference is much reduced in the right panel, where the masses of the simulated clusters are computed by applying (12). The reason for this difference between "true" and "recovered" masses is partly due to the violation of hydrostatic equilibrium, associated to subsonic gas bulk motions (e.g., Rasia et al. 2004; Nagai et al. 2007b) and partly to the poor fit provided by the β-model (e.g., Ascasibar et al. 2003) when extended to large radii. It is also interesting to note that the mass estimator of (12) under-predicts the intrinsic scatter of the M–T relation from the simulations. This is due to the fact that the equation

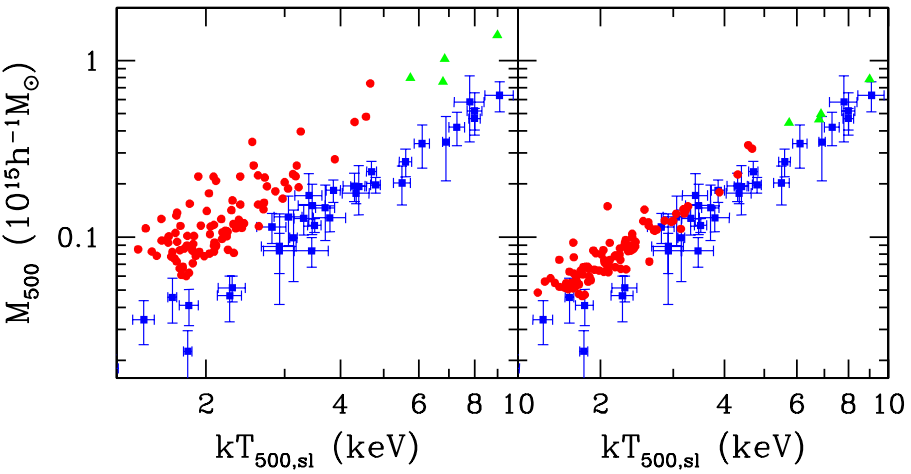


Fig. 7 The relation between mass and spectroscopic-like temperature within r_{500}. *Red circles* and *green triangles* are for simulations, which include cooling, star formation and feedback from galactic winds, while *squares with error bars* are the observational data by Finoguenov et al. (2001). *Left panel*: M_{500} exactly computed by summing the mass of all the particles within r_{500}. *Right panel*: M_{500} estimated as in the observational data, by using the equation of hydrostatic equilibrium for a polytropic β-model. From Rasia et al. (2005)

of hydrostatic equilibrium imposes a strong correlation between ICM temperature and total mass. Any scatter is then associated to a cluster-by-cluster variation of the parameters β and γ, which may not be fully representative of the diversity of the ICM structure among different objects.

Using XMM-Newton and Chandra data, different authors (e.g., Arnaud et al. 2005; Vikhlinin et al. 2005) applied the equation of hydrostatic equilibrium by avoiding the assumption of a simple beta-model for the gas density profile. As a result, the observed and the simulated M–T relations turned out to agree with each other, especially when simulated clusters are analysed in the same way as real clusters (e.g., Rasia et al. 2006; Nagai et al. 2007b). The left panel of Fig. 8 shows the comparison between the observational results by Arnaud et al. (2005) and simulations. Here the discrepancy with respect to the non-radiative runs by Evrard et al. (1996) is alleviated when including the effect of star formation and SN feedback (Borgani et al. 2004). In a similar way, the right panel shows the comparison between the observed M–T relation by Vikhlinin et al. (2005) and the simulations by Nagai et al. (2007a) who computed cluster masses by using the same method applied to the Chandra data. This plot further demonstrates the good level of agreement between simulated and observed M–T relation, once the 10–15 percent violation of hydrostatic equilibrium is taken into account.

3.3 The Mass–"Pressure" Relation

To first approximation, the ICM can be represented by a smooth gas distribution in pressure equilibrium within the cluster potential well. Therefore, the expectation is that gas pressure should be the quantity that is more directly correlated to the total collapsed mass. For this reason, any pressure-related observational quantity should provide a robust minimum-scatter proxy to the cluster mass. One such observable is the Comptonisation parameter, measured through the Sunyaev-Zeldovich effect (SZE; e.g., Carlstrom et al. 2002, for a review), which

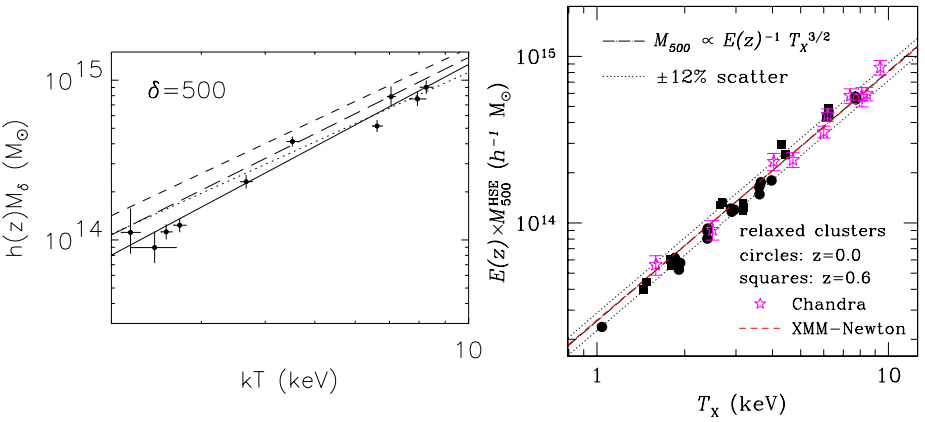


Fig. 8 *Left panel*: the observed mass–temperature relation at r_{500} for nearby clusters observed with XMM-Newton, compared with M–T relations from simulations (from Arnaud et al. 2005). The *solid line* is the best fit to observations, while the *dotted line* is the best fit only to a subsample of hot clusters. The *short-dashed line* is the relation from the non-radiative simulations by Evrard et al. (1996), while the *long-dashed line* is the relation from the radiative simulations by Borgani et al. (2004). *Right panel*: the M–T relation at r_{500} from the radiative simulations by Nagai et al. (2007a); *circles*: $z = 0$; *squares*: $z = 0.6$, which include cooling, star-formation and an inefficient form of SN feedback, compared to Chandra observations for a set of nearby relaxed clusters by Vikhlinin et al. (2005); *stars with error bars*. Masses of simulated clusters have been computed by using the same estimator, based on the hydrostatic equilibrium, applied to the Chandra data. Masses of both simulated and observed clusters have been rescaled with redshift according to the evolution predicted by the self-similar model

is proportional to the ICM pressure integrated along the line of sight. Observations of the SZ effect are now reaching a high enough quality for extended sets of clusters, to allow correlating the SZ signal with X-ray observable quantities. For instance, Bonamente et al. (2007) analysed a set of 38 massive clusters in the redshift range $0.14 \leq z \leq 0.89$, for which both SZE imaging from the OVRO/BIMA interferometric array and Chandra X-ray observations are available. As a result, they found that the slope and the evolution of the scalings of the Comptonisation parameter with gas mass, total mass and X-ray temperature are all in agreement with the prediction of the self-similar model.

In an attempt to provide an X-ray observable related to the pressure, Kravtsov et al. (2006) introduced the quantity $Y_X = M_{\rm gas} T$, defined by the product of the total gas mass times the temperature, both measured within a given aperture. With this definition, Y_X represents the X-ray counterpart of the Compton-y parameter, measured from the SZ effect. By computing this quantity for a set of simulated clusters, Kravtsov et al. (2006) showed that Y_X has a very tight correlation with the cluster mass, with a remarkably small scatter of only 8 per cent. The application to Chandra observations of a set of nearby relaxed clusters also shows that this relation has a comparably small scatter. Again, once masses of simulated and observed clusters are computed by applying the same hydrostatic estimator, the normalisation of the Y_X–M relation from models and data closely agree with each other (Nagai et al. 2007a; see left panel of Fig. 9). Since mass is determined in this case from temperature, Y_X and M are not independent quantities and, therefore, the scatter in their scaling relation may be underestimated.

Maughan (2007) computed Y_X for an extended set of clusters extracted from the Chandra archive, in the redshift range $0.1 < z < 1.3$. The results of his analysis, shown in the right panel of Fig. 9, indicate that Y_X is also tightly correlated with the X-ray luminosity through

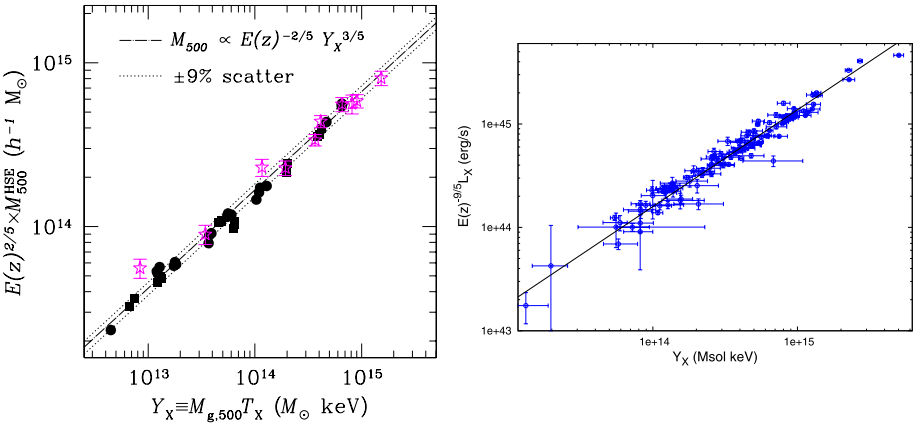


Fig. 9 *Left panel*: the relation between $Y_X = M_{\rm gas} T_X$ and M_{500}. *Points with error bars* are for Chandra observational data (from Nagai et al. 2007a). Masses of simulated clusters have been computed by using the same estimator, based on the hydrostatic equilibrium, applied to the Chandra data. Masses of both simulated and observed clusters have been rescaled with redshift according to the evolution predicted by the self-similar model. *Right panel*: the relation between Y_X and the X-ray luminosity, both estimated within r_{500}, for a set of clusters at $0.1 < z < 1.3$ extracted from the Chandra archive (from Maughan 2007). Luminosities have been rescaled according to the redshift dependence expected from self-similar evolution

a relation which evolves in a self-similar way. Again, since the gas mass is obtained from the X-ray surface brightness, Y_X and L_X are not independent quantities, thus possibly leading to an underestimate of the intrinsic scatter in their scaling relation.

3.4 The Gas Mass Fraction

The measurement of the baryon mass fraction in nearby galaxy clusters has been recognised for several years to be a powerful method to measure the cosmological density parameter (e.g., White et al. 1993; Mohr et al. 1999), while its redshift evolution provides constraints on the dark energy content of the Universe (e.g., Allen et al. 2002; Ettori et al. 2003, and references therein). Since diffuse gas dominates the baryon budget of clusters, a precise measurement of the ICM total mass represents a fundamental step in the application of this cosmological test. While this method relies on the basic assumption that all clusters contain baryons in a cosmic proportion, a number of observational evidences show that the gas mass fraction is smaller in lower temperature systems (Lin et al. 2003; Sanderson et al. 2003). This fact forces one to restrict the application to the most massive and relaxed systems. In addition, since X-ray measurements of the gas mass fraction are generally available only out to a fraction of the cluster virial radius, the question then arises as to whether the gas fraction in these regions is representative of the cosmic value. Indeed, observational evidence has been found for an increase of the gas mass fraction with radius (e.g., Castillo-Morales and Schindler 2003).

In this respect, hydrodynamical simulations offer a way to check how the gas mass is distributed within individual clusters and as a function of the cluster mass, thus possibly providing a correction for such biases. Indeed, Allen et al. (2004) resorted to the set of SPH clusters simulated by Eke et al. (1998) to calibrate the correction factor that one needs to apply to extrapolate the baryon fraction computed at r_{2500}, which is the typical radius at which it is measured, to the cosmic value. However, since this set of simulations does not

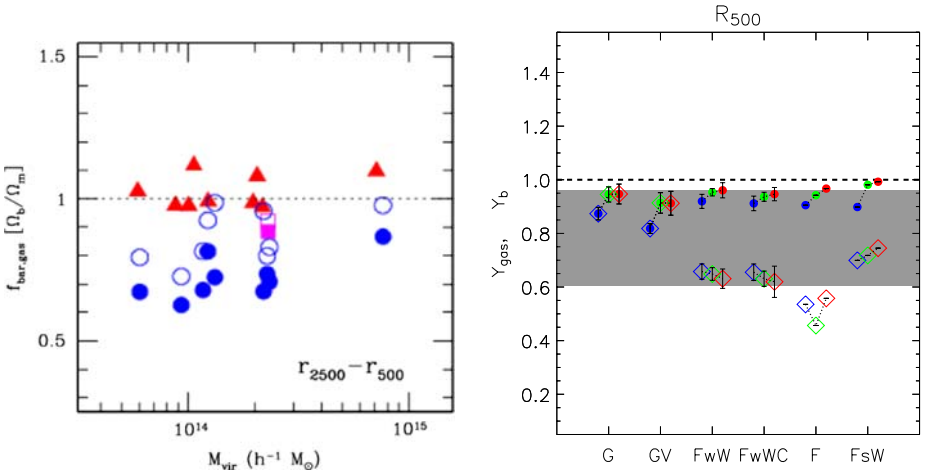


Fig. 10 *Left panel*: the gas fraction (*filled circles*) and the baryon fraction (*open circles*) for cluster simulations which include cooling and star formation. *Squares* indicate runs where cooling is switched off at $z < 2$, while *triangles* are for non-radiative simulations (from Kravtsov et al. 2005). *Right panel*: the gas fraction (*diamonds*) and the baryon fraction (*circles*) within r_{500} for SPH simulations of clusters, which include different gas physics. G: gravitational heating only; GV: G with a scheme for reduced SPH gas viscosity; FwW: radiative runs with feedback and weak galactic winds; FwWC: like FwW with also thermal conduction, with an efficiency of one-third of the Spitzer value; F: radiative runs with no winds; FsW: radiative runs with strong winds. *Blue*, *green* and *red symbols* refer to redshifts $z = 0$, 0.7 and 1 (from Ettori et al. 2006). The *shaded area* indicates the range of values of the observed gas fraction (Ettori and Fabian 1999)

include the effects of cooling, star formation and feedback heating, it is not clear whether they provide a reliable description of the gas distribution within real clusters.

Kravtsov et al. (2005) used high resolution simulations, based on an Eulerian code, for a set of clusters using both non-radiative and radiative physics. They found that including cooling and star formation has a substantial effect on the total baryon fraction in the central cluster regions, where it is even larger than the cosmic value. As shown in the left panel of Fig. 10, at the virial radius the effect is actually inverted, with the baryon fraction of radiative runs lying below that of the non-radiative runs. In addition, they also compared results obtained from Eulerian and SPH codes and found small, but systematic, differences between the resulting baryon fractions.

Ettori et al. (2006) performed a similar test, based on SPH simulations, but focusing on the effect of changing in a number of ways the description of the relevant physical processes, such as gas viscosity, feedback strength and thermal conduction. The results of their analysis, which is shown in the right panel of Fig. 10, demonstrated that the baryon fraction is generally stable but only at rather large radii, $\gtrsim r_{500}$. They also showed that changing the description of the relevant ICM physical processes changes the extrapolation of the baryon fraction from the central regions, relevant for X-ray measurements, while also slightly affecting the redshift evolution.

On the one hand, these results show that simulations can be used as calibration instruments for cosmological applications of galaxy clusters. On the other hand, they also demonstrate that for this calibration to reach the precision required to constrain the dark energy content of the Universe, one needs to include in simulations the relevant physical processes which determine the ICM observational properties.

4 Profiles of X-ray Observables

4.1 The Temperature Profiles

Already ASCA observations, despite their modest spatial resolution, have established that most of the clusters show significant departures from isothermality, with negative temperature gradients characterised by a remarkable degree of similarity, out to the largest radii sampled (e.g., Markevitch et al. 1998). Besides confirming the presence of these gradients, Beppo-SAX observations (e.g., De Grandi 2002) showed that they do not extend down to the innermost cluster central regions, where instead an isothermal regime is observed, possibly followed by a decline of the temperature towards the centre, at least for relaxed clusters. The much improved sensitivity of the Chandra satellite provides now a more detailed picture of the central temperature profiles (e.g., Vikhlinin et al. 2005; Baldi et al. 2007). At the same time, a number of analyses of XMM–Newton observations now consistently show the presence of a negative gradient at radii $\gtrsim 0.1 r_{200}$ (e.g., Piffaretti et al. 2005; Pratt et al. 2007, and references therein). Relaxed clusters are generally shown to have a smoothly declining profile toward the centre, reaching values which are about half of the overall virial cluster temperature in the innermost sampled regions, with non-relaxed clusters having, instead, a larger variety of temperature profiles. The emerging picture suggests that gas cooling is responsible for the decline of the temperature in the central regions, while some mechanism of energy feedback should be responsible for preventing overcooling, thereby suppressing the mass deposition rate and the resulting star formation.

As for hydrodynamical simulations, they have shown to be generally rather successful in reproducing the declining temperature profiles outside the core regions (e.g., Loken et al. 2002; Roncarelli et al. 2006), where gas cooling is relatively unimportant. On the other hand, as we have already discussed in Sect. 2.2, including gas cooling has the effect of steepening the T-profiles in the core regions, in clear disagreement with observations. The problem of the central temperature profiles in radiative simulations has been consistently found by several independent analyses (e.g., Valdarnini 2003; Borgani et al. 2004; Nagai et al. 2007a) and is interpreted as due to the difficulty that currently implemented feedback schemes have in balancing the cooling runaway.

As an example, we show in the left panel of Fig. 11 the comparison between simulated and observed temperature profiles, recently presented by Nagai et al. (2007a), which is based on a set of clusters simulated with an Adaptive Mesh Refinement (AMR) code. This plot clearly shows that the central profiles of simulated clusters are far steeper than the observed ones, by an amount which increases when cooling and star formation are turned on. Although this result is in qualitative agreement with other results based on SPH codes, an eye-ball comparison with the left panel of Fig. 3 shows a significant difference of the profiles in the central regions for the non-radiative runs. Indeed, while SPH simulations generally show a flattening at $r \lesssim 0.1 r_{\rm vir}$, Eulerian simulations are instead characterised by continuously rising profiles. While this difference is reduced when cooling is turned on, it clearly calls for the need of performing detailed comparisons between different simulation codes, in the spirit of the Santa Barbara Cluster Comparison Project (e.g., Frenk et al. 1999), and to understand in detail the reason for these differences.

As discussed above, the steep temperature profiles predicted in the central regions witness the presence of overcooling. Viceversa, the fact that a feedback mechanism is able to produce the correct temperature profiles does not guarantee in itself that overcooling is prevented. Indeed, Kay et al. (2007) found that their "targeted" scheme of SN feedback produces temperature profiles which are in reasonable agreement with observations (see the

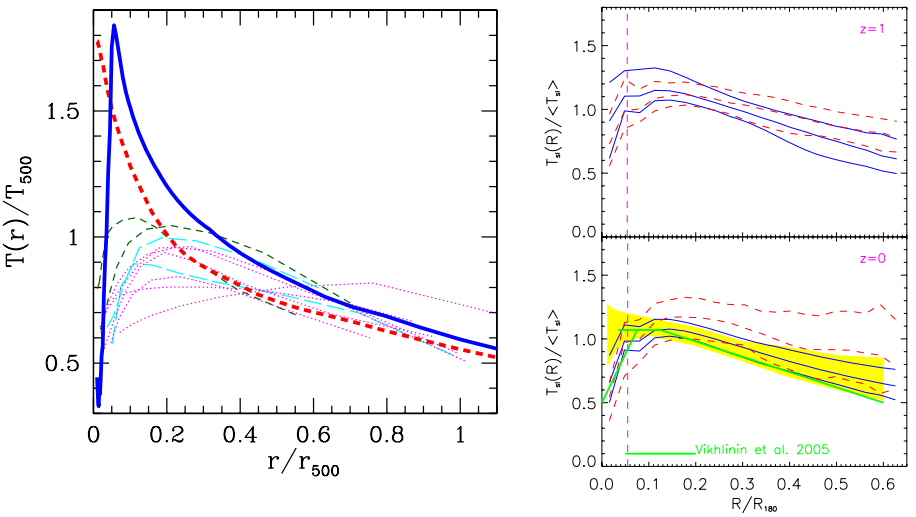


Fig. 11 *Left panel*: the temperature profiles in AMR simulations of clusters (*thick curves*) and in real clusters (*thin curves*). The simulation results are the average over 16 numerical clusters. The *solid line* is for runs with cooling and star formation, while the *dashed line* is for non-radiative simulations. The observational curves are for clusters with different temperatures (from Nagai et al. 2007a). *Right panel*: Temperature profiles at $z = 1$ (*top*) and at $z = 0$ (*bottom*) from a set of SPH simulated clusters including cooling, star-formation and a "targeted" scheme of SN feedback (from Kay et al. 2007). Median and 10/90 percentiles are shown. *Solid* and *dashed lines* are for irregular and regular clusters, respectively. The *vertical dashed line* indicates the smallest scale which is numerically resolved. In the *bottom panel*, the *heavy green line* shows the profile from Chandra observations of cool core clusters (Vikhlinin et al. 2005), while the *yellow shaded area* is for a XMM-Newton sample of nearby clusters (Pratt et al. 2007)

right panel of Fig. 11). However, even with this efficient feedback scheme, the resulting stellar fraction within clusters was still found to be too high, $\gtrsim 25$ per cent, thus indicating the presence of a substantial overcooling in their simulations.

Clearly, resolving the discrepancy between observed and simulated central temperature profiles requires that simulations are able to produce the correct thermal structure of the observed "cool cores". This means that a suitable feedback should compensate the radiative losses of the gas at the cluster centre, while keeping it at about $\sim 1/3$ of the virial temperature. A number of analyses converge to indicate that AGN should represent the natural solution to this problem. Considerable efforts have been spent to investigate how cooling can be self-consistently regulated by feedback from a central AGN in static cluster potentials (e.g., Churazov et al. 2001; Omma et al. 2004; Brighenti and Mathews 2006; Sternberg et al. 2007, and reference therein), while only quite recently these studies have been extended to clusters forming in a cosmological context (e.g., Heinz et al. 2006; Sijacki et al. 2007). Although the results of these analyses are quite promising, we still lack for a detailed comparison between observational data and an extended set of cosmological simulations of clusters, convincingly showing that AGN feedback is able to provide the correct ICM thermal structure for objects spanning a wide range of masses, from poor groups to rich massive clusters.

4.2 The Entropy Profiles

As already mentioned in Sect. 2, a convenient way of characterising the thermodynamical properties of the ICM is through the entropy, which, in X-ray cluster studies, is usually defined as $S = T/n_e^{2/3}$ (see the review by Voit 2005, for a detailed discussion about the role of entropy in cluster studies). Since the current numerical description of the heating/cooling interplay in the central cluster regions is unable to reproduce the observed temperature structure, there is no surprise that simulations have difficulties also in accounting for the observed entropy structure. However, while the ICM thermodynamics is sensitive to complex physical processes in the core regions, one expects simulations to fare much better in the outer regions, say at $r > 0.2\,r_{200}$, where the gas dynamics should be dominated by gravitational processes. This expectation is also supported by the agreement between the observed and the simulated slope of the temperature profiles outside cluster cores.

If gravity were the only process at work, then the prediction of the self-similar model is that $S \propto T$. Therefore, by plotting profiles of the reduced entropy, S/T, one expects them to fall on the top of each other for clusters of different temperatures. However, observations revealed that this is not the case. Indications from ASCA data (Ponman et al. 2003) showed that entropy profiles of poor clusters and groups have an amplitude which is higher than that expected for rich clusters from the above scaling argument. This result has been subsequently confirmed by the XMM-Newton data analysed by Pratt and Arnaud (2005) and by Piffaretti et al. (2005). Quite remarkably, a relatively higher entropy for groups is found not only in the central regions, where it can arise as a consequence of the heating/cooling processes, but extends to all radii sampled by X-ray observations, out to $\lesssim 0.5 r_{500}$. These analyses consistently indicate that $S \propto T^\alpha$, with $\alpha \simeq 0.65$, instead of $\alpha = 1$ as expected from self-similar scaling.

Ponman et al. (2003) and Voit et al. (2003) interpreted this entropy excess in poor clusters and groups as the effect of entropy amplification generated by shocks from smoothed gas accretion. The underlying idea is the following. In the hierarchical scenario for structure formation, a galaxy group is expected to accrete from relatively smaller filaments and merging sub-groups than a rich cluster does. Suppose now to heat the gas with a fixed amount of specific energy (or entropy). Such a diffuse heating will be more effective to smooth the accretion pattern of a group than that of a rich cluster, just as a consequence of the lower virial temperature of the structures falling into the former object. In this case, accretion shocks take place at a lower density and, therefore, are more efficient in generating entropy. While predictions from the semi-analytical approach by Voit et al. (2003) are in reasonable agreement with observational results, one may wonder whether these predictions are confirmed by full hydrodynamical simulations.

To tackle this problem, Borgani et al. (2005) performed a series of SPH hydrodynamical simulations of four galaxy clusters and groups, with temperature in the range 0.5–3 keV, using both non-radiative and radiative runs, also exploring a variety of heating recipes. As a result, they found that galactic ejecta powered by an even extremely efficient SN feedback are not able to generate the observed level of entropy amplification. This result is shown in the left panel of Fig. 12 where the profiles of reduced entropy are plotted for the simulated structures in the case of moderate (upper panel) and very strong (bottom panel) galactic outflows. Borgani et al. also showed that only adding an entropy floor at $z = 3$ provides an efficient smoothing of the gas accretion pattern and, therefore, a substantial degree of entropy amplifications (see Fig. 4). Although this diffuse pre-heating provides an adequate entropy amplification at large radii, the price to pay is that the entropy level is substantially

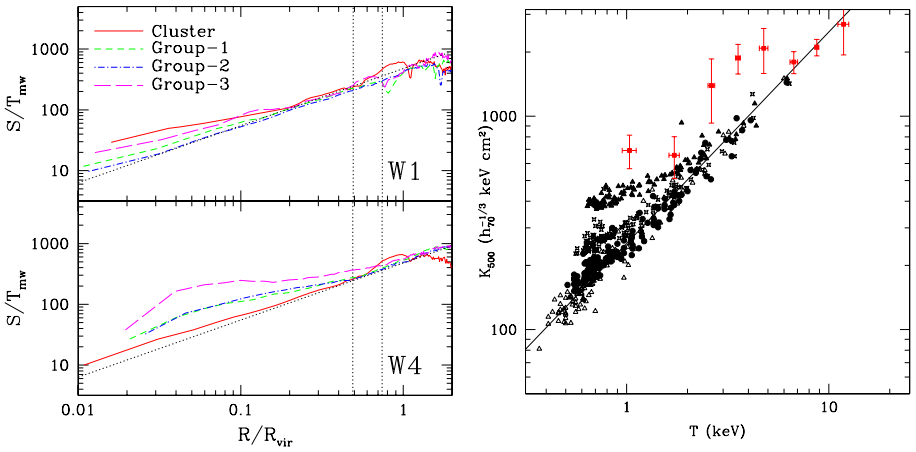


Fig. 12 *Left panel*: the profile of reduced entropy, S/T, for a series of SPH GADGET runs of one cluster and three groups, including cooling, star formation and feedback from galactic winds. The *top panel* is for runs in which the winds have a velocity of about 360 km s^{-1}, while the *bottom panel* is for an extreme wind model with a velocity of 830 km s^{-1} (from Borgani et al. 2005). The *straight dotted line* marks the slope $S \propto R^{0.95}$ which is the best-fitting to the observed entropy profiles (Pratt and Arnaud 2005; Piffaretti et al. 2005). *Right panel*: the relation between entropy computed at r_{500} for clusters and groups identified in AMR ENZO cosmological simulations (from Younger and Bryan 2007), compared with the observational results by Ponman et al. (2003); *symbols with error bars*. Shown are four different cases of entropy injection at $z = 10$: no pre-heating (*open triangles*), 78 keV cm^2 (*filled circles*), 155 keV cm^2 (*stars*), 311 keV cm^2 (*filled triangles*). The *solid line* is a power-law fit to the self-similar prediction from the simulations

increased in the central regions. This is at variance with high-resolution Chandra measurements of low entropy gas in the innermost cluster regions, where it reaches values as low as ~ 10 keV cm^2 (Donahue et al. 2006).

A similar analysis has also been performed by Younger and Bryan (2007), who used a set of AMR non-radiative cosmological simulations performed with the ENZO code (O'Shea et al. 2004), with gas entropy boosted at high redshift, $z = 10$. In keeping with previous analyses, they found that pre-heated simulations are generally able to reproduce the observed luminosity–temperature and mass–temperature relations. However, differently from Borgani et al. (2005), even the most extreme pre-heating scheme does not provide an appreciable degree of entropy amplification. This result is shown in the right panel of Fig. 12. The deviations from self-similarity for the entropy computed at r_{500} are always too small to reconcile the simulations with the observational data by Ponman et al. (2003). A possible reason for the different results with respect to the analysis by Borgani et al. is due to the higher redshift of pre-heating, $z = 10$ instead of 3, used by Younger & Bryan. Furthermore, grid-based codes are known to produce entropy profiles that, in the central part of the halos, $r \lesssim 0.2 r_{200}$, are flatter than those produced by SPH codes (e.g., Voit et al. 2005). Therefore, the question arises as to whether different numerical schemes of hydrodynamics react in different ways to pre-heating. It would be highly recommendable that future simulation comparison projects will include tests of how different codes behave in the presence of simple schemes of non-gravitational heating.

5 Summary

In this paper we reviewed the current status of the comparison between cosmological hydrodynamical simulations of the X-ray properties of the intra-cluster medium (ICM) and observations. We first presented the basic predictions of the self-similar model, based on the assumption that only gravitational processes drive the evolution of the ICM. We then showed how a number of observational facts are at variance with these predictions, with poor clusters and groups characterised by a relatively lower density and higher entropy of the gas. This calls for the need of introducing some extra physical processes, such as non-gravitational heating and radiative cooling, which are able to break the self-similarity between objects of different size. The results based on simulations, which include these effects, can be summarised as follows.

(1) The observed scaling relation between X-ray luminosity and temperature can be reproduced by simulations including cooling only, but at the expense of producing a too large fraction of cooled gas. Introducing ad-hoc schemes of entropy injection at high redshift (the so-called pre-heating) can eventually produce the correct L_X–T relation. However, no simulations have been presented so far, in which a good agreement with observations is achieved by using a feedback scheme in which the energy release and thermalisation from SN or AGN is self-consistently computed by the simulated star formation or accretion onto a cosmologically evolving population of black holes.

(2) Simulations which include cooling, star formation and SN feedback, correctly reproduce the observed mass–temperature and "pressure"–temperature relations. Quite interestingly, this agreement is achieved once the masses of simulated clusters are estimated by applying the same mass estimators, based on the assumption of hydrostatic equilibrium, which are applied to the analysis of real clusters. The violation of hydrostatic equilibrium, related to the non-thermal pressure support from subsonic gas motions, would otherwise contribute to a ~ 20 per cent overestimate of the normalisation of the above scaling relations for simulated clusters.

(3) Both non-radiative and radiative runs naturally predict the observed negative gradients of the temperature profiles outside the cluster core regions, $r \gtrsim 0.2 r_{200}$. This suggests that the ICM thermal structure in these regions is indeed dominated by the action of gravitational processes, such as heating from shocks associated to supersonic gas accretion.

(4) Introducing cooling has the effect of steepening the temperature profiles in the innermost regions, as a consequence of the adiabatic compression of inflowing gas, caused by the lack of pressure support. Therefore, gas in the cores of simulated clusters generally lies in a high-temperature phase, which is very well separated from the cold ($\sim 10^4$ K) phase. This is at variance with the observed "cool core" structure of real clusters, which show instead the presence of a fair amount of gas, down to a limiting temperature of $\sim 1/3$ of the virial temperature, which formally has a rather short cooling time. The presence of this gas makes the observed temperature profiles of relaxed clusters to peak at $r \simeq 0.2 r_{200}$, thereby decreasing by about a factor two in the innermost sampled regions. The discrepancy between the simulated and the observed "cool core" calls for the need of introducing in cluster simulations a suitable feedback mechanism which is able to compensate the radiative losses, keeps pressurised gas in the central regions, and suppresses the mass deposition rate.

(5) Simulations are generally rather successful in reproducing the observed slope of the entropy profiles outside the core regions, $S \propto r^\alpha$ with $\alpha \simeq 1$. This slope is also close to that predicted by models based on gravitational heating from spherical accretion, thus lending further support to the picture that gravity drives the evolution of the ICM outside the core regions. Quite intriguingly, however, the normalisation of the profiles out to the largest sampled radii, $\lesssim r_{500}$, is observed to have a milder temperature dependence than expected from

the self-similar model, $S \propto T^{\alpha}$ with $\alpha \simeq 0.65$ instead of $\alpha \simeq 1$. This entropy excess in poor systems can be reproduced in simulations only by adding a diffuse pre-heating mechanism, which smoothes the pattern of gas accretion, thereby leading to an amplification of entropy generation from shocks in poorer systems. However, this mechanism is able to provide the correct explanation only for clusters with temperature $T \lesssim 3$ keV. Generating entropy amplification in hotter systems would require an exceedingly large amount of pre-heating, which would generate too shallow entropy profiles.

In general, the above results demonstrate the capability of cosmological hydrodynamical simulations to predict the correct thermodynamical properties of the ICM outside the central regions of galaxy clusters. On the other hand, even simulations based on an efficient SN feedback fail in preventing overcooling and providing the correct description of the thermodynamical ICM properties in the central regions. The "cool core" failure of simulations based on stellar feedback is also witnessed by the exceedingly massive and blue central galaxies that are generally produced (see also Borgani et al. 2008—Chap. 18, this issue).

The fact that cool cores are observed for systems spanning at least two orders of magnitude in mass, from $\sim 10^{13} M_{\odot}$ to $\sim 10^{15} M_{\odot}$, suggests that only one self-regulated mechanism should be mainly responsible for this, rather than the combination of several mechanisms possibly acting over different time-scales. Although AGN are generally thought to naturally provide such a mechanism, a detailed comparison between observational data and an extended set of cluster simulations, including this kind of feedback is still lacking. Furthermore, understanding in detail the role of AGN in determining the ICM properties requires addressing in detail two major issues. Firstly, the accretion process onto the central black hole involves scales of the order of the parsec, while the X-ray observable effects involve scales of $\sim$ 10–100 kpc. This requires understanding the cross-talk between two ranges of scales, which differ by at least four orders of magnitude. Secondly, the total energy budget available from AGN is orders of magnitude larger than that required to regulate cooling flows. Therefore, one needs ultimately to understand the channels for the thermalisation of the released energy and how this naturally takes place without resorting to any ad-hoc tuning. These problems definitely need to be addressed by simulations of the next generation, which will be aimed at understanding in detail the evolution of the cosmic baryons and the observational X-ray properties of galaxy clusters.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialised X-ray components in clusters of galaxies". SB wishes to thank Hans Böhringer, Stefano Ettori, Gus Evrard, Alexey Finoguenov, Andrey Kravtsov, Pasquale Mazzotta, Silvano Molendi, Giuseppe Murante, Rocco Piffaretti, Trevor Ponman, Elena Rasia, Luca Tornatore, Paolo Tozzi, Alexey Vikhlinin and Mark Voit for a number of enlightening discussions. Partial support from the PRIN2006 grant "Costituenti fondamentali dell'Universo" of the Italian Ministry of University and Scientific Research and from the INFN grant PD51 is also gratefully acknowledged.

References

S.W. Allen, A.C. Fabian, Mon. Not. R. Astron. Soc. **297**, L57 (1998)
S.W. Allen, R.W. Schmidt, A.C. Fabian, Mon. Not. R. Astron. Soc. **334**, L11 (2002)
S.W. Allen, R.W. Schmidt, H. Ebeling, A.C. Fabian, L. van Speybroeck, Mon. Not. R. Astron. Soc. **353**, 457 (2004)
N. Arimoto, Y. Yoshii, Astron. Astrophys. **173**, 23 (1987)
M. Arnaud, A.E. Evrard, Mon. Not. R. Astron. Soc. **305**, 631 (1999)
M. Arnaud, N. Aghanim, D.M. Neumann, Astron. Astrophys. **389**, 1 (2002)
M. Arnaud, E. Pointecouteau, G.W. Pratt, Astron. Astrophys. **441**, 893 (2005)
Y. Ascasibar, G. Yepes, V. Müller, S. Gottlöber, Mon. Not. R. Astron. Soc. **346**, 731 (2003)

A. Baldi, S. Ettori, P. Mazzotta, P. Tozzi, S. Borgani, Astrophys. J. **666**, 835 (2007)
M.L. Balogh, F.R. Pearce, R.G. Bower, S.T. Kay, Mon. Not. R. Astron. Soc. **326**, 1228 (2001)
J.J. Bialek, A.E. Evrard, J.J. Mohr, Astrophys. J. **555**, 597 (2001)
M. Bonamente, M. Joy, S. LaRoque et al., Astrophys. J. (2007, in press). arXiv:astro-ph/0708.815
S. Borgani, in Lectures for 2005 Guillermo Haro Summer School on Clusters. Lect. Notes Phys. (Springer) (2006, in press). astro-ph/0605575
S. Borgani, F. Governato, J. Wadsley et al., Astrophys. J. **559**, L71 2001)
S. Borgani, F. Governato, J. Wadsley et al., Mon. Not. R. Astron. Soc. **336**, 409 (2002)
S. Borgani, G. Murante, V. Springel et al., Mon. Not. R. Astron. Soc. **348**, 1078 (2004)
S. Borgani, A. Finoguenov, S.T. Kay et al., Mon. Not. R. Astron. Soc. **361**, 233 (2005)
S. Borgani, K. Dolag, G. Murante et al., Mon. Not. R. Astron. Soc. **367**, 1641 (2006)
S. Borgani, D. Fabjan, L. Tornatore et al., Space Sci. Rev. (2008). doi:10.1007/s.11214-008-9322-7
M. Branchesi, I.M. Gioia, C. Fanti, R. Fanti, Astron. Astrophys. **472**, 739 (2007)
F. Brighenti, W.G. Mathews, Astrophys. J. **643**, 120 (2006) (2006)
G.L. Bryan, M.L. Norman, Astrophys. J. **495**, 80 (1998)
A. Bykov, K. Dolag, F. Durret, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9312-9
J.E. Carlstrom, G.P. Holder, E.D. Reese, Anu. Rev. Astron. Astrophys. **40**, 643 (2002)
A. Castillo-Morales, S. Schindler, Astron. Astrophys. **403**, 433 (2003)
A. Cavaliere, R. Fusco-Femiano, Astron. Astrophys. **49**, 137 (1976)
E. Churazov, M. Brüggen, C.R. Kaiser, H. Böhringer, W. Forman, Astrophys. J. **554**, 261 (2001)
R. Davé, N. Katz, D.H. Weinberg, Astrophys. J. **579**, 23 (2002)
S. De Grandi, Molendi, S., Astrophys. J. **567**, 163 (2002)
K. Dolag, A. Evrard, M. Bartelmann, Astron. Astrophys. **369**, 36 (2001)
K. Dolag, M. Jubelgas, V. Springel, S. Borgani, E. Rasia, Astrophys. J. **606**, L97 (2004)
K. Dolag, S. Borgani, S. Schindler, A. Diaferio, A. Bykov, Space Sci. Rev. (2008a). doi:10.1007/s11214-008-9316-5
K. Dolag, A. Bykov, A. Diaferio, Space Sci. Rev. (2008b). doi:10.1007/s11214-008-9319-2
M. Donahue, D.J. Horner, K.W. Cavagnolo, G.M. Voit, Astrophys. J. **643**, 730 (2006)
V.R. Eke, J.F. Navarro, C.S. Frenk, Astrophys. J. **503**, 569 (1998)
S. Ettori, A.C. Fabian, Mon. Not. R. Astron. Soc. **305**, 834 (1999)
S. Ettori, P. Tozzi, P. Rosati, Astron. Astrophys. **398**, 879 (2003)
S. Ettori, S. Borgani, L. Moscardini et al., Mon. Not. R. Astron. Soc. **354**, 111 (2004a)
S. Ettori, P. Tozzi, S. Borgani, P. Rosati, Astron. Astrophys. **417**, 13 (2004b)
S. Ettori, K. Dolag, S. Borgani, G. Murante, Mon. Not. R. Astron. Soc. **365**, 1021 (2006)
A.E. Evrard, J.P. Henry, Astrophys. J. **383**, 95 (1991)
A.E. Evrard, C.A. Metzler, J.F. Navarro, Astrophys. J. **469**, 494 (1996)
A. Finoguenov, T.H. Reiprich, H. Böhringer, Astron. Astrophys. **368**, 749 (2001)
C.S. Frenk, S.D.M. White, P. Bode et al., Astrophys. J. **525**, 554 (1999)
S. Heinz, M. Brüggen, A. Young, E. Levesque, Mon. Not. R. Astron. Soc. **373**, L65 (2006)
S.F. Helsdon, T.J. Ponman, Mon. Not. R. Astron. Soc. **315**, 356 (2000)
J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-y
N. Kaiser, Mon. Not. R. Astron. Soc. **222**, 323 (1986)
N. Kaiser, Astrophys. J. **383**, 104 (1991)
N. Katz, S.D.M. White, Astrophys. J. **412**, 455 (1993)
S.T. Kay, A.C. da Silva, N. Aghanim et al., Mon. Not. R. Astron. Soc. **377**, 317 (2007)
O. Kotov, A. Vikhlinin, Astrophys. J. **633**, 781 (2005)
A.V. Kravtsov, D. Nagai, A.A. Vikhlinin, Astrophys. J. **625**, 588 (2005)
A.V. Kravtsov, A. Vikhlinin, D. Nagai, Astrophys. J. **650**, 128 (2006)
G.F. Lewis, A. Babul, N. Katz et al., Astrophys. J. **536**, 623 (2000)
Y.-T. Lin, J.J. Mohr, S.A. Stanford, Astrophys. J. **591**, 749 (2003)
C. Loken, M.L. Norman, E. Nelson et al., Astrophys. J. **579**, 571 (2002)
D.H. Lumb, J.G. Bartlett, A.K. Romer et al., Astron. Astrophys. **420**, 853 (2004)
M. Markevitch, Astrophys. J. **504**, 27 (1998)
M. Markevitch, W.R. Forman, C.L. Sarazin, A. Vikhlinin, Astrophys. J. **503**, 77 (1998)
B.J. Maughan, Astrophys. J. **668**, 772 (2007)
B.J. Maughan, L.R. Jones, H. Ebeling, C. Scharf, Mon. Not. R. Astron. Soc. **365**, 509 2006)
P. Mazzotta, E. Rasia, L. Moscardini, G. Tormen, Mon. Not. R. Astron. Soc. **354**, 10 (2004)
J.J. Mohr, B. Mathiesen, A.E. Evrard, Astrophys. J. **517**, 627 (1999)
O. Muanwong, P.A. Thomas, S.T. Kay, F.R. Pearce, H.M.P. Couchman, Astrophys. J. **552**, L27 (2001)
O. Muanwong, P.A. Thomas, S.T. Kay, F.R. Pearce, Mon. Not. R. Astron. Soc. **336**, 527 (2002)

O. Muanwong, S.T. Kay, P.A. Thomas, Astrophys. J. **649**, 640 (2006)
D. Nagai, A.V. Kravtsov, A. Vikhlinin, Astrophys. J. **668**, 1 (2007a)
D. Nagai, A. Vikhlinin, A.V. Kravtsov, Astrophys. J. **655**, 98 (2007b)
J.F. Navarro, C.S. Frenk, S.D.M. White, Mon. Not. R. Astron. Soc. **275**, 720 (1995)
J. Nevalainen, M. Markevitch, W. Forman, Astrophys. J. **532**, 694 (2000)
H. Omma, J. Binney, G. Bryan, A. Slyz, Mon. Not. R. Astron. Soc. **348**, 1105 (2004)
B.W. O'Shea, G. Bryan, J. Bordner et al., in Adaptive mesh refinement – theory and applications. Springer Lect. Notes Comp. Sci. Eng. (2004)
J.P.F. Osmond, T.J. Ponman, Mon. Not. R. Astron. Soc. **350**, 1511 (2004)
C. Pfrommer, T.A. Enßlin, V. Springel, M. Jubelgas, K. Dolag, Mon. Not. R. Astron. Soc. **378**, 385 (2007)
R. Piffaretti, P. Jetzer, J.S. Kaastra, T. Tamura, Astron. Astrophys. **433**, 101 (2005)
T.J. Ponman, D.B. Cannon, J.F. Navarro, Nature **397**, 135 (1999)
T.J. Ponman, A.J.R. Sanderson, A. Finoguenov, Mon. Not. R. Astron. Soc. **343**, 331 (2003)
G.W. Pratt, M. Arnaud, Astron. Astrophys. **429**, 791 (2005)
G.W. Pratt, H. Böhringer, J.H. Croston et al., Astron. Astrophys. **461**, 71 (2007)
E. Rasia, G. Tormen, L. Moscardini, Mon. Not. R. Astron. Soc. **351**, 237 (2004)
E. Rasia, P. Mazzotta, S. Borgani et al., Astrophys. J. **618**, L1 (2005)
E. Rasia, S. Ettori, L. Moscardini et al., Mon. Not. R. Astron. Soc. **369**, 2013 (2006)
M. Roncarelli, S. Ettori, K. Dolag et al., Mon. Not. R. Astron. Soc. **373**, 1339 (2006)
P. Rosati, S. Borgani, C. Norman, Anu. Rev. Astron. Astrophys. **40**, 539 (2002)
A.J.R. Sanderson, T.J. Ponman, A. Finoguenov, E.J. Lloyd-Davies, M. Markevitch, Mon. Not. R. Astron. Soc. **340**, 989 (2003)
C.L. Sarazin, *X-Ray Emission from Clusters of Galaxies*. Cambridge Astrophys. Ser. (Cambridge Univ. Press, Cambridge, 1988)
S. Schindler, Astron. Astrophys. **305**, 756 (1996)
D. Sijacki, V. Springel, T. Di Matteo, L. Hernquist, Mon. Not. R. Astron. Soc. **380**, 877 (2007)
V. Springel, L. Hernquist, Mon. Not. R. Astron. Soc. **339**, 289 (2003a)
Springel, V., Hernquist, L., Mon. Not. R. Astron. Soc. **339**, 312 (2003b)
A. Sternberg, F. Pizzolato, N. Soker, Astrophys. J. **656**, L5 (2007)
R.S. Sutherland, M.A. Dopita, Astrophys. J. **88**, 253 (1993)
L. Tornatore, S. Borgani, V. Springel et al., Mon. Not. R. Astron. Soc. **342**, 1025 (2003)
P. Tozzi, C. Norman, Astrophys. J. **546**, 63 (2001)
R. Valdarnini, Mon. Not. R. Astron. Soc. **339**, 1117 (2003)
A. Vikhlinin, Astrophys. J. **640**, 710 (2006)
A. Vikhlinin, M. Markevitch, S.S. Murray et al., Astrophys. J. **628**, 655 (2005)
G.M. Voit, Rev. Mod. Phys. **77**, 207 (2005)
G.M. Voit, G.L. Bryan, Nature **414**, 425 (2001)
G.M. Voit, M.L. Balogh, R.G. Bower, C.G. Lacey, G.L. Bryan, Astrophys. J. **593**, 272 (2003)
G.M. Voit, S.T. Kay, G.L. Bryan, Mon. Not. R. Astron. Soc. **364**, 909 (2005)
S.D.M. White, J.F. Navarro, A.E. Evrard, C.S. Frenk, Nature **366**, 429 (1993)
D.A. White, C. Jones, W. Forman, Mon. Not. R. Astron. Soc. **292**, 419 (1997)
X.-P. Wu, Y.-J. Xue, Astrophys. J. **572**, L19 (2002)
J.D. Younger, G.L. Bryan, Astrophys. J. **666**, 647 (2007)

Chapter 14
Numerical Simulations of the Warm-Hot Intergalactic Medium

Serena Bertone · Joop Schaye · Klaus Dolag

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9318-3

Abstract In this paper we review the current predictions of numerical simulations for the origin and observability of the warm hot intergalactic medium (WHIM), the diffuse gas that contains up to 50 per cent of the baryons at $z \sim 0$. During structure formation, gravitational accretion shocks emerging from collapsing regions gradually heat the intergalactic medium (IGM) to temperatures in the range $T \sim 10^5$–10^7 K. The WHIM is predicted to radiate most of its energy in the ultraviolet (UV) and X-ray bands and to contribute a significant fraction of the soft X-ray background emission. While O VI and C IV absorption systems arising in the cooler fraction of the WHIM with $T \sim 10^5$–$10^{5.5}$ K are seen in *FUSE* and *Hubble Space Telescope* observations, models agree that current X-ray telescopes such as *Chandra* and *XMM-Newton* do not have enough sensitivity to detect the hotter WHIM. However, future missions such as *Constellation-X* and *XEUS* might be able to detect both emission lines and absorption systems from highly ionised atoms such as O VII, O VIII and Fe XVII.

Keywords Large-scale structure of the universe · Intergalactic medium · Diffuse radiation · X-rays: diffuse background · Quasars: absorption lines

S. Bertone (✉)
Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK
e-mail: serena@scipp.ucsc.edu

Present address:
S. Bertone
Santa Cruz Institute for Particle Physics, University of California at Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

J. Schaye
Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands
e-mail: schaye@strw.leidenuniv.nl

K. Dolag
Max-Planck-Institut für Astrophysik, P.O. Box 1317, 85741 Garching, Germany
e-mail: kdolag@mpa-garching.mpg.de

1 Introduction

The cosmic baryon abundance was first inferred by applying the theory of primordial nucleosynthesis to observations of light element abundances (e.g., Olive et al. 2000) and has been confirmed by observations of the cosmic microwave background radiation (CMB). A recent measurement from the Wilkinson Microwave Anisotropy Probe reveals that $\Omega_{\rm b}h^2 = 0.0223^{+0.0007}_{-0.0009}$, where the Hubble parameter is $h = 0.73^{+0.03}_{-0.04}$ (Spergel et al. 2007). At $z > 2$ most of the baryons are thought to reside in the diffuse, photoionised intergalactic medium (IGM) with $T \sim 10^4$–10^5 K, traced by the Lyα forest[1] (e.g., Rauch et al. 1997; Weinberg et al. 1997; Schaye 2001). The gas traced by the Lyα forest, together with the mass in galaxies, fully accounts for the cosmic baryon abundance at $z \sim 2$ and confirms the prediction of nucleosynthesis theories and the CMB measurements.

However, when the mass of stars, interstellar gas and plasma in clusters of galaxies at redshift $z \sim 0$ is considered, only a small fraction of the mass budget can be accounted for (e.g., Persic and Salucci 1992; Fukugita et al. 1998; Fukugita and Peebles 2004). While the gas traced by Lyα forest may account for about a third of the low-z baryons (e.g., Danforth and Shull 2007), the majority of the baryons remain invisible. Unless the baryon budget at high redshift has been overestimated by two different measurements, a large amount of the low redshift baryons must be "missing".

One of the basic predictions of cosmological, gas-dynamical simulations is the distribution of baryons in the universe. How many baryons are locked in stars? How many reside in clusters and filaments? What is the state of the diffuse gas? Simulations of structure formation predict that the missing baryons problem is measure of our technological limitations rather than a real problem. In other words, the baryons are out there, but we cannot see them because they reside in gas that is too dilute to be detected in emission and too hot to be traced by Lyα absorption.

The idea of gravitational heating of the intergalactic gas was first suggested by Sunyaev and Zel'dovich (1972) and subsequently developed (e.g., Nath and Silk 2001; Furlanetto and Loeb 2004; Rasera and Teyssier 2006). While the largest structures in the universe form, the IGM is heated by gravitational shocks that efficiently propagate from the collapsing regions to the surrounding medium. Simulations predict that gas compressed and heated by shocks can reach temperatures of 10^8 K in rich clusters of galaxies, while filaments and mildly overdense regions are heated to temperatures in the range 10^5 to 10^7 K (e.g., Cen and Ostriker 1999; Davé et al. 2001; Cen and Ostriker 2006). The latter are commonly known as the Warm-Hot Intergalactic Medium (WHIM), which represents about half the total baryonic mass in the universe at $z \sim 0$.

At $T \sim 10^5$–10^7 K, the IGM is collisionally ionised and becomes transparent to Lyα radiation. As a consequence, the H I Lyα forest does not trace the bulk of the gas mass at low redshift. Although the shock-heated IGM emits radiation in the UV and in the soft X-ray bands, the total energy radiated away by mildly overdense gas is orders of magnitude too small to be detected by current instruments. Similarly, the absorption from such a gas along the line of sight to a bright X-ray source is too weak to be resolved by current spectrographs (Richter et al. 2008—Chap. 3, this issue).

[1] Since the diffuse intergalactic medium is highly ionised, the baryon density inferred from Lyα forest observations is inversely proportional to the assumed ionisation rate. Since independent constraints on the UV intensity are only accurate to a factor of few at best, we cannot prove that most of the baryons reside in the forest.

In this paper, we review the predictions of numerical simulations for the properties and the observability of the WHIM. In Sect. 2 we briefly mention the numerical techniques used to simulate the IGM and we discuss the mechanisms that produce the WHIM, namely heating by gravitational accretion shocks and other, non-gravitational, heating processes. A more detailed description of the numerical techniques themselves can be found in Dolag et al. (2008—Chap. 12, this issue). Section 3 describes the contribution of the WHIM emission to the soft X-ray background. The current predictions for the detectability of the WHIM in emission and absorption are discussed in Sect. 4 and in Sect. 5 respectively. Section 6 briefly reports on the effect of the WHIM on the CMB radiation and Sect. 7 discusses the magnetisation of the IGM. We draw our conclusions in Sect. 8.

2 Origin and Properties of the WHIM

2.1 Numerical Techniques

In this subsection, we briefly discuss the numerical techniques that have been used to investigate the IGM and the WHIM in particular. However, we refer the reader to the review of Dolag et al. (2008—Chap. 12, this issue) for a thorough description of specific simulation techniques).

The most common approach to model the WHIM is through cosmological hydrodynamical simulations. The strength of hydrodynamical simulations is that the basic gas dynamical processes that determine the evolution of the diffuse gas are directly simulated and that it is possible to make virtual observations of the WHIM with existing or future instruments. However, *ad hoc* empirical prescriptions are needed to include processes such as galactic winds and feedback from active galactic nuclei (AGN), whose physics is still poorly understood and in any case would require a resolution far beyond the current standards to be included self-consistently. Hydrodynamical simulations of the WHIM have typically used either Smoothed Particle Hydrodynamics (SPH, e.g., Hellsten et al. 1998; Croft et al. 2001; Davé et al. 2001; Yoshikawa et al. 2004; Yoshikawa and Sasaki 2006; Dolag et al. 2006), a uniform grid (e.g., Cen and Ostriker 1999, 2006), or Adaptive Mesh Refinement (AMR, e.g., Kravtsov et al. 2002) to solve the gas dynamics.

Several properties of the WHIM, including the impact of gravitational and non-gravitational heating processes on the IGM, and the observability of emission lines and absorption systems in UV and X-ray spectra, have been successfully investigated with analytical and semi-analytical techniques (Perna and Loeb 1998; Pen 1999; Nath and Silk 2001; Valageas et al. 2002; Viel et al. 2003; Furlanetto et al. 2004; Viel et al. 2005; Furlanetto et al. 2005; Rasera and Teyssier 2006).

2.2 Gravitational Heating

Theoretical predictions and numerical simulations agree that the heating of the IGM is fuelled by the gravitational accretion shocks emerging during the collapse of cosmic structures (e.g., Cen and Ostriker 1999; Davé et al. 2001; Nath and Silk 2001; Valageas et al. 2002; Kang et al. 2005; Cen and Ostriker 2006; Bykov et al. 2008—Chap. 7, this issue). Simulations suggest that the distribution of the warm-hot gas closely reflects the distribution of the shock waves, and therefore the underlying density field. The role of shock waves in heating the IGM has been investigated using cosmological hydrodynamical simulations by Kang et al. (2005), who find that the fastest shocks can be found around clusters and groups of

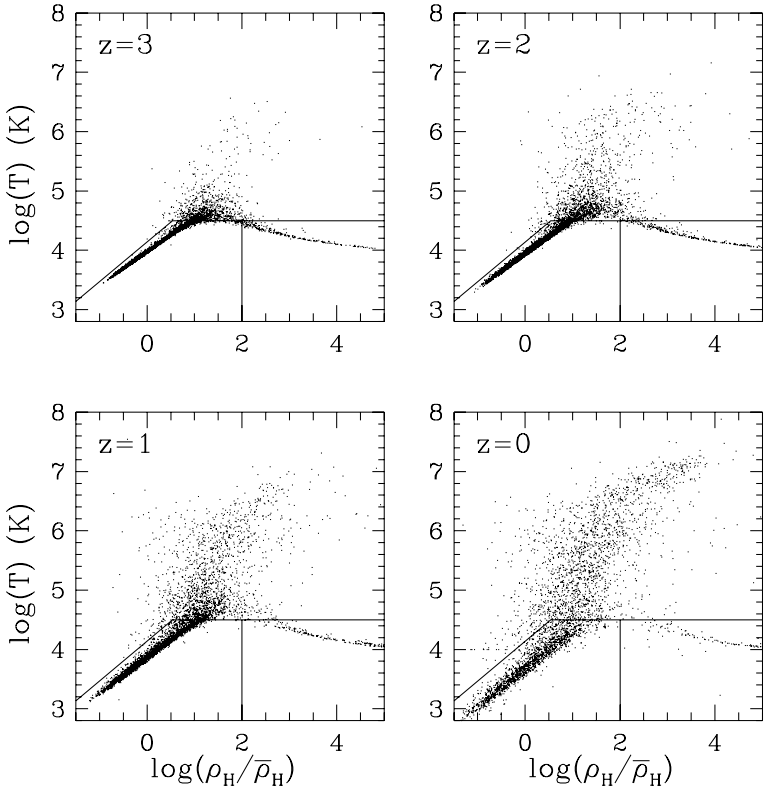


Fig. 1 The temperature vs. overdensity plane at $z = 3$, 2, 1, and 0. The gas mostly resides in three phases: a cool, photoionised, diffuse IGM (*tail to lower left*); the warm-hot shocked IGM; and gas condensed into stars and star forming clouds (*tail to lower right*). The *solid lines* indicate an approximate separation between these phases. From Davé et al. (1999)

galaxies, while the slowest shocks usually propagate across low density regions like sheets and filaments. This reflects the fact that the infall velocity is a measure of the depth of the potential well. Furlanetto et al. (2005) reach similar conclusions using analytical techniques and find that the shocks with the highest temperatures are associated with the most massive virialised objects, up to distances larger than their virial radii.

The WHIM gas is organised in a complex structure of filaments and sheets, that generally reflects the dynamics of structure formation. According to Dolag et al. (2006), the gas density in filaments is typically 10–100 times the mean density and varies more smoothly than in clusters. Filaments have a coherence length of about 5 Mpc, although some can be as long as 25 Mpc, and a diameter of about 3–5 Mpc. The detailed thermal structure of the WHIM strongly depends on the direction of propagation of the accretion shocks, while the velocity field appears to be orthogonal to the filaments at large distances from galaxy clusters and aligned to the filaments at small distances.

Most simulations predict that between 30 and 50 per cent of all the baryons in the low redshift universe have temperatures in the interval 10^5–10^7 K (e.g., Hellsten et al. 1998; Cen and Ostriker 1999; Davé et al. 1999, 2001; Cen and Ostriker 2006). Figure 1 shows the distribution of the gas in the temperature-density plane at $z = 3$, 2, 1, and 0. Hot gas with $T > 10^7$ K resides in the hottest cores of groups and in large clusters of galaxies.

Most of the gas, however, is at temperatures $T < 10^7$ K and at $z = 0$ is about evenly spread between the WHIM and the cooler, mostly photo-ionised IGM. At higher redshifts, a larger fraction of the baryons resides in the photo-ionised IGM. Kang et al. (2005) find that, besides the gas shock-heated to $T = 10^5$–10^7 K, a significant amount of mass, distributed mostly along sheet-like structures, is shock-heated to $T < 10^5$ K by shock waves propagating with velocities $v_{\rm sh} < 150\,{\rm km\,s^{-1}}$. This mostly photo-ionised gas is responsible for the majority of the O V and O VI absorption lines, but contributes little to the total O V and O VI emission, which is instead produced by warmer, collisionally ionised ions.

2.3 Non-Equilibrium Calculations

Simulations that include non-equilibrium processes have been performed by Teyssier et al. (1998), Cen and Fang (2006) and Yoshikawa and Sasaki (2006). In most hydrodynamical simulations, it is routinely assumed that diffuse gas is in ionisation equilibrium (Cen and Ostriker 1999; Davé et al. 2001; Yoshikawa et al. 2003, 2004; Kang et al. 2005; Roncarelli et al. 2006). In general, ionisation equilibrium is nearly achieved in the centres of galaxy clusters and in photo-ionised regions with $T < 10^5$ K. However, this may not happen in the outer regions of clusters, in groups, and in the WHIM (Yoshikawa et al. 2003).

Similar conclusions are reached by Cen and Fang (2006) and Yoshikawa and Sasaki (2006). Cen and Fang (2006) claim that observational results for the abundance of O VI absorption lines are better reproduced by models including non-equilibrium processes. Yoshikawa and Sasaki (2006) find that a significant fraction of the WHIM at $z = 0$ is not in ionisation equilibrium, as do Yoshida et al. (2005). However, Yoshikawa and Sasaki (2006) calculate that although this strongly affects the relative abundance of ions such as O VI, O VII and O VIII, it does not have a significant effect on the observable signatures of the WHIM in emission or absorption.

When gas is infalling onto a collapsing structure and is heated by gravitational accretion shocks, most of its energy is carried by ions, while electrons can only gain thermal energy through collisions with ions. The temperatures of ions and electrons can therefore be very different and converge only if energy can be efficiently transferred from ions to electrons. If only Coulomb interactions are considered, then the relaxation timescale can be shorter than the age of the universe for the WHIM (Yoshikawa et al. 2003; Yoshida et al. 2005). Since the relative abundance of metal ions and the plasma emissivity are sensitive to the electron temperature, this would have consequences for the observability of the WHIM.

Yoshida et al. (2005) introduce a two-temperature model for the thermal evolution of an ionised gas. Their model includes an explicit treatment for collisional relaxation processes between electrons and ions and allows one to follow the temperature of free electrons separately from the temperature of ions. They find that at $z = 0$ most WHIM gas shows a two-temperature structure. Figure 2 shows the electron temperature in the two-temperature model compared to the mean temperature of the gas at $z = 0$ for the simulations of Yoshida et al. (2005). Clearly, the electron temperature of the WHIM is significantly lower than the mean gas temperature, when non-equilibrium processes are taken into account. Yoshida et al. (2005) calculate that the relative abundance of O VI, O VII and O VIII predicted by a two-temperature model can be about an order of magnitude different from that predicted by a one-temperature model. This yields similarly different values for the line emission of ions.

However, from supernova remnants we know that plasma waves are typically much more efficient than collisional processes in transferring energy from ions to electrons (e.g., Rakowski 2005) and these effects have not yet been included in simulations of the WHIM. Yoshida et al. (2005) may therefore have overestimated the importance of the two-temperature structure of the post-shock gas.

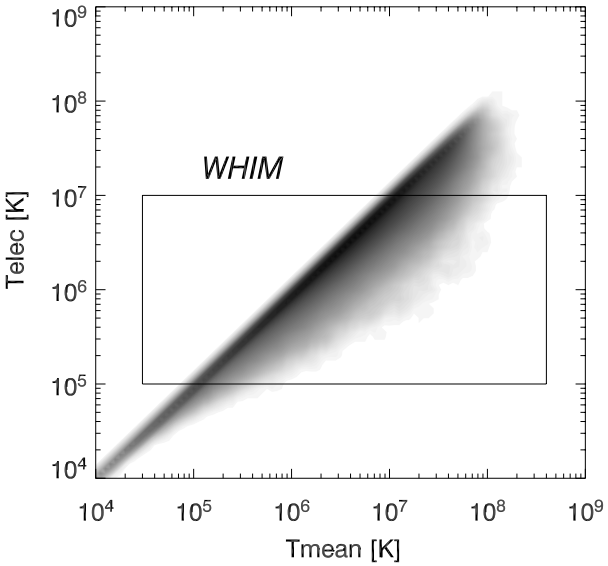


Fig. 2 Electron temperature versus the mean gas temperature. The *grey scale* shows the number density of gas particles and the *rectangular box* encloses the WHIM gas, with $10^5 < T_{\rm elec} < 10^7$ K. From Yoshida et al. (2005)

2.4 Non-Gravitational Physics

Several studies have investigated whether sources of non-gravitational heating might significantly affect the thermal state of the IGM. Simulations by Cen and Ostriker (2006) predict that non-gravitational processes, such as galactic winds and X-ray emission from galaxies and quasars, contribute no more than 20 per cent of the energy required to heat the IGM, in agreement with previous results from Davé et al. (2001) and Nath and Silk (2001).

A powerful test to constrain the contribution of non-gravitational processes is the comparison of the observed intensity of the soft X-ray background to the predictions of numerical simulations. Since most of the observed background emission has been resolved into point sources (Hasinger et al. 1998; Worsley et al. 2005), the comparison of the predictions of hydrodynamical simulations with observations can put tight constraints on the amount of the observed background emission that can be attributed to the diffuse intergalactic gas.

In general, the intensity of the soft X-ray background predicted by simulations that include only gravitational heating processes strongly exceeds the observed emission from high density, collapsed regions (e.g., Nath and Silk 2001; Bryan and Voit 2001; Davé et al. 2001; Cen and Ostriker 2006). The emission is primarily due to Bremsstrahlung and reflects the cooling of gas in groups and clusters of galaxies. Analytical calculations by Pen (1999) and Wu et al. (2001) find an order of magnitude more background light than predicted by simulations, owing to the importance of gravitational pre-heating. They suggest that additional heating from non-gravitational processes might be necessary to "puff up" and hence decrease the emission from collapsed regions and reconcile the theoretical predictions with the observations. However, they note that non-gravitational heating does not seem to be relevant for uncollapsed regions like filaments and voids.

Voit and Bryan (2001) argue that non-gravitational processes can help to establish an entropy floor in high density regions. The lowest entropy material will drop out and form stars unless it is heated through feedback from star formation. They suggest that an entropy floor of $\sim$100 keV cm^{-2} is required to reconcile the theoretical results with observations. Bryan and Voit (2001) demonstrate further that radiative cooling and additional heating from

feedback can efficiently suppress the X-ray background, while substantially modifying the characteristic spectral shape of the background itself.

3 WHIM Contribution to the Soft X-Ray Background

The extended spatial distribution over cosmological scales makes the WHIM an essential contributor to the diffuse X-ray background radiation. The most intense X-ray emission comes from relatively compact sources such as clusters, active galactic nuclei and galaxies. The gas in these objects has high emissivity power, but accounts for only a small fraction of the baryon mass in the universe. The emissivity of the WHIM is usually a few orders of magnitude lower, because of its lower density and temperature. However, the WHIM accounts for up to half the baryonic mass at $z \sim 0$ and its integrated emission over cosmological scales might be roughly comparable to that of dense regions and compact objects.

Analytical calculations (Valageas et al. 2002) and numerical simulations predict that the WHIM contributes between a few per cent (Kang et al. 2005) and 40 per cent (Croft et al. 2001; Roncarelli et al. 2006) of the total X-ray background emission, consistent with the observational upper limit of $(1.2 \pm 0.3) \times 10^{-12}$ erg s^{-1} cm^{-2} deg^{-2} (Worsley et al. 2005). Because of the softness of the spectrum, the WHIM produces only between 4 per cent (Croft et al. 2001) and 10 per cent (Roncarelli et al. 2006) of the total hard X-ray emission. According to the simulations of Roncarelli et al. (2006), about 90 per cent of the X-ray background emission is produced at $z < 0.9$. Figure 3 shows simulated maps of the soft (0.5–2 keV, upper panels) and hard (2–10 keV, lower panels) X-ray emission from a region 3.78 degree on a side. The left panels show the total emission from the gas, while the right panels show the contribution of the WHIM.

Phillips et al. (2001) estimate that in hydrodynamical simulations the WHIM gas at $z = 0$ is responsible for about 5–15 per cent of the total extragalactic surface brightness in the energy range 0.5–2 keV (Fig. 4), corresponding to a flux of 0.24×10^{-12} erg s^{-1} cm^{-2} deg^{-2}. The WHIM line emission peaks in the 0.5–0.8 keV energy range, where observations by *ASCA* and *ROSAT* show a spectral bump, as shown in Fig. 4. Phillips et al. (2001) argue that the bump is due to WHIM emission, and in particular to metal line emission. Similarly, Kravtsov et al. (2002) produce simulations that faithfully reconstruct the local universe and find that about 5–10 per cent of the X-ray background at energies of about 1 keV may be contributed by the WHIM. Ursino and Galeazzi (2006) find a contribution of up to 20 per cent in the energy range 0.37–0.925 keV, most of which is emitted by filaments at redshifts between 0.1 and 0.6. Dense and bright filaments are present in about 10 per cent of a generic field of view, and can account for more than 20 per cent of the total emission. In such cases, redshifted emission lines from highly ionised atoms should be clearly detectable.

3.1 Soft X-Ray Excess Radiation in Galaxy Clusters

Soft X-ray excess radiation in clusters has been detected in observations by the Extreme Ultraviolet Explorer (*EUVE*), *ROSAT* and *XMM-Newton* (Lieu et al. 1996; Bonamente et al. 2001; Finoguenov et al. 2003; see Durret et al. 2008—Chap. 4, this issue for an overview). The temperature of the ICM is usually estimated by assuming a one-temperature plasma model for the diffuse gas. The soft X-ray flux calculated from the observed X-ray emission of clusters, however, often appears to be above the expected thermal contribution. It has been suggested that the excess radiation may be due to foreground emission from the WHIM, or from warm diffuse gas in the clusters.

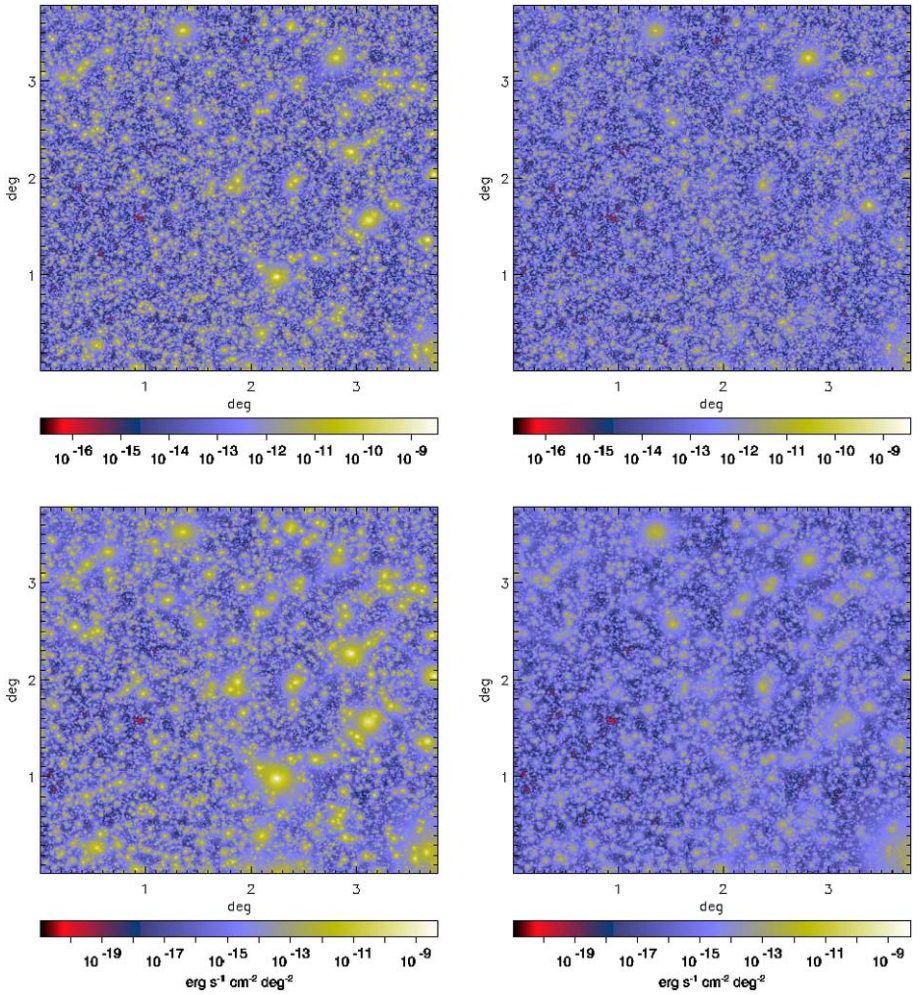


Fig. 3 Maps of the soft (0.5–2 keV, *upper panels*) and hard (2–10 keV, *lower panels*) X-ray intensity obtained by considering all the gas particles (*left panels*), and only the gas particles with temperatures in the range $10^5 < T < 10^7$ K (WHIM, *right panels*). The maps are 3.78 deg on a side, with a pixel size $(1.66 \text{ arcsec})^2$. The two maps show results for the same realisation of the past light-cone. From Roncarelli et al. (2006)

Cheng et al. (2005) investigate the origin of this excess with a set of cluster simulations. They find that the presence of WHIM is not necessary to explain the observations, because the excess emission is produced by extremely dense gas associated with merging substructures within the cluster. In clusters where this gas is present, the temperature of the inner regions is usually 30–50 per cent above that predicted by the one-temperature plasma model. Dolag et al. (2006) find that, since the WHIM surface brightness in X-rays reaches at most 10^{-16} erg s^{-1} cm^{-2} arcmin^{-2}, projection effects of WHIM filaments in front of galaxy clusters can account for no more than 10 per cent of the cluster emission. This may partly explain the soft X-ray excess, but clearly does not fully account for it. Similar conclusions have been reached by Mittaz et al. (2004).

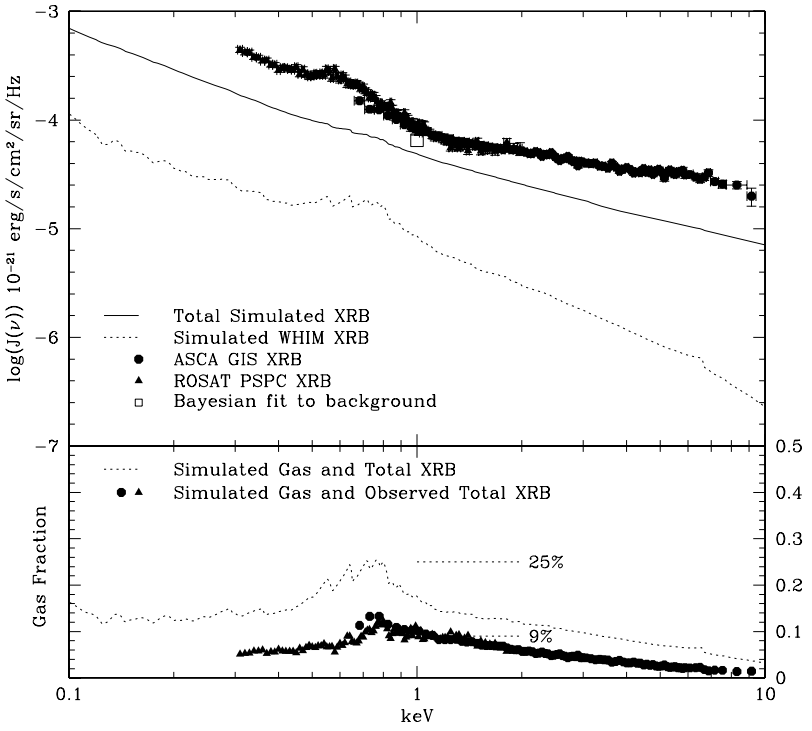


Fig. 4 The predicted integrated X-ray background (*solid curve*) and the predicted contributions from the IGM (*dashed curve*) and WHIM alone (*dotted curve*). The observations of the XRB spectrum from the ASCA LSS region (Miyaji et al. 2000) and ROSAT PSPC fields are shown along with the Bayesian fit to observations of the X-ray background (Barcons et al. 2000). The fractional contribution of the WHIM to the total simulated and observed backgrounds is plotted in the *lower panel*. Phillips et al. (2001) argue that the factor of two difference between the normalisation of the simulated background and that of the observed background may be due to the low value of Ω_b assumed in the simulations or to the underestimate of the observed AGN formation rate used to compute the spectrum. From Phillips et al. (2001)

4 Line Emission

Most of the energy radiated by the WHIM in the UV and soft X-ray bands is emitted through emission lines. The detection of emission lines from the WHIM is key to understanding the properties of the gas itself. The relative intensities of emission lines from highly ionised metals in the WHIM gives information about the temperature of the emitting gas, while the absolute intensities reflect the density and the degree of metal enrichment of the gas.

The relative abundances of ionised atoms is determined by the temperature and, if photoionisation plays a role, the density of the WHIM. Ions whose abundance peaks at the highest temperatures, such as Fe XXV and Fe XXVI, are concentrated in the hottest, highest density regions, while ions that peak at lower temperatures, such as O VII, O VIII and Fe XVII, are more widespread and originate in groups and filaments (Fang et al. 2002). The detection of WHIM emission from the intracluster medium of filaments, groups and cluster with different masses helps to understand the impact of feedback in these structures (Pierre et al. 2000).

4.1 Soft X-Rays

Yoshikawa et al. (2003) and Yoshikawa et al. (2004) find that the temperature of the IGM can be estimated if both the O VII and O VIII ions are simultaneously observed. However, if only one line is detectable, then the line cannot be identified unambiguously. Figure 5 shows template spectra of a collisionally ionised plasma with temperature $T = 10^6$ K, $10^{6.5}$ K, and 10^7 K, and with metallicity $Z = Z_\odot$ and $Z = 0.1Z_\odot$, which display the contribution of oxygen ions as a function of temperature. Clearly, the detectability of the line emission depends strongly on the metallicity.

It is still controversial if the WHIM can actually be detected in emission by current and future X-ray telescopes. Voit et al. (2001) calculate that the virialised regions of groups and clusters cover about a third of the sky. This is a potential problem in the identification of X-ray emission from the WHIM in filaments and distant groups, because it creates source-confusion and strongly contaminates the observations. Pierre et al. (2000) predict that *XMM-Newton* should be able to detect diffuse emission from strong filaments at $z \sim 0.1$, although no such detections have been confirmed so far. As Fang et al. (2005) calculate, X-ray emission lines from the WHIM are hard to detect in high density regions and in voids. The emission from clusters and groups is dominated by the hot intracluster medium, while in voids the X-ray spectrum is completely dominated by radiation from background AGN and

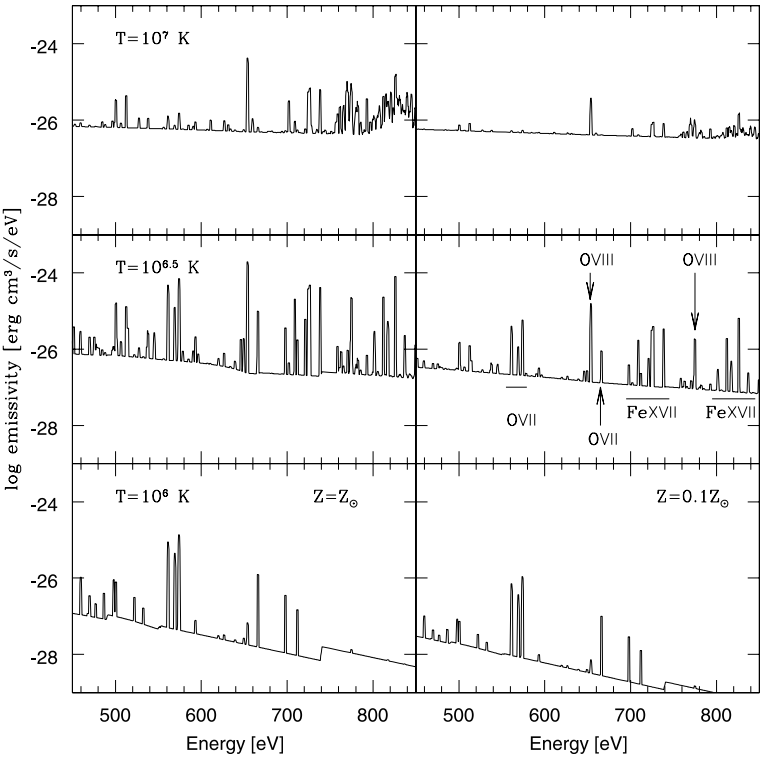


Fig. 5 Template spectra of a collisionally ionised plasma with temperature $T = 10^6$ K (*lower panels*), $10^{6.5}$ K (*middle panels*), and 10^7 K (*upper panels*). Spectra for metallicity $Z = Z_\odot$ and $Z = 0.1Z_\odot$ are shown in the *left* and *right panels*, respectively. From Yoshikawa et al. (2003)

by the Galactic foreground. In both cases, the emission of the WHIM would be too low to be distinguished from the dominant component. On the other hand, it might be possible to resolve at least a few strong emission lines from the WHIM in filamentary regions, despite the fact that most of the emission is contributed by AGN and by the Galactic foreground.

Fang et al. (2005) argue that planned telescopes such as *Constellation-X* and the X-ray Evolving Universe Spectroscopy mission (*XEUS*) have little chance to see a significant signal for the WHIM emission, because their effective collecting area is too small to detect diffuse line emission from an extended area. Ideally, a mission to detect line emission from the WHIM should have a large field of view, coupled with a high resolving power and spectral resolution. Yoshikawa et al. (2003) and Yoshikawa et al. (2004) use large scale simulations and simulations that reproduce the local universe to produce mock X-ray spectra of the WHIM emission. The simulated maps of the local universe (Yoshikawa et al. 2004) demonstrate that WHIM gas in filamentary structures can be successfully imaged in targeted observations by the proposed Diffuse Intergalactic Oxygen Surveyor (*DIOS*), an X-ray telescope of new generation which combines a large field of view with high spatial and spectral resolution. *DIOS* and several other proposed missions are discussed in more detail in Paerels et al. (2008—Chap. 19, this issue).

4.2 UV

While O VII and O VIII lines are emitted by gas with $T \gtrsim 10^6$ K, line emission from the O VI $\lambda\lambda$1032, 1038 Å and C IV $\lambda\lambda$1548, 1551 Å doublets traces the coolest fraction of the WHIM, with $T \sim 10^5$–$10^{5.5}$ K. For the bulk of the WHIM mass, the line strengths are predicted to be a few orders of magnitude lower than the background (Furlanetto et al. 2004). Kravtsov et al. (2002) estimate the intensity of the O VI emission line at $\delta \sim 10$ and column densities $N_{\mathrm{O\,VI}} \sim 10^{14}$ cm^{-2} to be $I \sim 1 - 10$ photons cm^{-2} s^{-1} sr^{-1}, which is an order of magnitude below the sensitivity of past and current telescopes such as the Far Ultraviolet Spectroscopic Explorer (*FUSE*) and *SPEAR*.

However, Furlanetto et al. (2004) predict that O VI and C IV line emission could be detected above the background from enriched dense regions with sizes of about 50–100 kpc. Furlanetto et al. (2004) argue that the connection between the line emission and the enrichment by galactic winds may help understand the chemical enrichment history of the IGM and the physics of winds.

5 Absorption

The X-ray forest is the high-energy counterpart of the Lyα forest. While the Lyα forest is produced by cool gas absorbing Lyα photons in the optical spectra of high redshift quasars, the X-ray forest is the result of absorption by highly ionised atoms in the spectra of bright X-ray sources (Richter et al. 2008—Chap. 3, this issue).

The detection of the X-ray forest in high resolution spectra of quasars and gamma ray bursts (GRB) is possibly one of the most powerful tools to investigate the properties of the WHIM in the low redshift universe. For example, the number density of absorbers gives an estimate of the baryon density times the metallicity of the IGM, while the ratio between the O VII and the O VIII line strengths gives information on the distribution of the gas temperature and density.

A thorough knowledge of the physical state of the WHIM, as Furlanetto et al. (2005) point out, helps to understand the mechanisms that have determined its evolution in a cosmological context. Furlanetto et al. (2005) use analytic techniques to investigate how the

accretion shocks that heat up the IGM during the formation of cosmic structures can give rise to absorption systems such as O VI and O VII. They find that the observed column densities of absorbers are predicted by post-shock cooling models, when both slow and fast cooling channels are taken into account. The observed number density of absorbers can be realistically reproduced if each virialised structure is surrounded by a network of shocks with a total cross section a few times the size of the virialised region. In this scenario, fast cooling would produce stronger O VII absorbers associated to O VI systems than expected from models of collisional ionisation equilibrium. O VII absorption systems should be the most common because the abundance of the O VII ion dominates over a large temperature range.

5.1 Soft X-Rays

The observability of the WHIM in the X-ray forest by current and future facilities has been a very popular topic in recent years (Richter et al. 2008—Chap. 3, this issue). A variety of predictions has been provided by hydrodynamical simulations (e.g., Hellsten et al. 1998; Kravtsov et al. 2002; Chen et al. 2003; Cen and Fang 2006; Kawahara et al. 2006) and semi-analytic models (e.g., Perna and Loeb 1998; Viel et al. 2003, 2005).

Kravtsov et al. (2002) find that the regions of the highest column densities of O VII and O VIII correspond to the high density regions in and around galaxy groups. The gas in these regions, with $T = 10^6$ K and $\delta \sim 100$, should produce the strongest O VII and O VIII absorption lines, with equivalent width $W > 100$ km s^{-1} (Hellsten et al. 1998). These lines should occur on average once per sight line in the redshift range $z = 0$–0.3 (Hellsten et al. 1998; Cen and Fang 2006). Cen and Fang (2006) predict abundances for O VI, O VII, and O VIII lines in the range 50–100 per unit redshift at $W = 1$ km s^{-1}, decreasing to 10–20 per unit redshift at $W = 10$ km s^{-1}.

Chen et al. (2003) suggest that the most promising strategy to find O VII and O VIII absorbers in the X-ray forest would be to search at the redshifts of known O VI absorbers. However, since O VI absorption traces gas at somewhat lower temperatures than O VII and O VIII, only future missions like *Constellation-X* and *XEUS* will be able to find absorption systems at higher temperatures.

Most predictions agree that *Chandra* and *XMM-Newton* do not have enough sensitivity to detect the WHIM in absorption. Chen et al. (2003) claim that a few strong O VII and O VIII absorption systems might be within reach of *Chandra* and *XMM-Newton*, but the probability to detect such lines in random lines of sight is less than 5 per cent (Kravtsov et al. 2002).

Future telescopes such as *Constellation-X* and *XEUS* might have a better chance to detect the WHIM in absorption, thanks to higher spectral resolution and larger collecting areas. *Constellation-X* should have a 50 per cent probability to detect the strongest O VII and O VIII absorption lines (Kravtsov et al. 2002). Because of its high sensitivity at high energies, it should also be able to detect absorption from ions such as Ne IX and Fe XVII, whose intensity peaks at about the same temperature as O VII, and Fe XX lines, which appear for $T \gtrsim 10^7$ K (Hellsten et al. 1998). Perna and Loeb (1998) predict that one day of integration time will allow *Constellation-X* and *XEUS* to detect a few O VIII absorption lines per unit redshift in the spectra of X-ray bright quasars. According to Viel et al. (2003), *Constellation-X* will detect about 6 absorption lines per unit redshift with $W > 10$ km s^{-1}, while *XEUS* could detect up to 30 lines with $W > 1$ km s^{-1} per unit redshift. Under the assumption of constant IGM metallicity, Kawahara et al. (2006) estimate that *XEUS* will be able to detect on average two O VII absorption systems with $\sigma > 3$ along any line of sight to bright quasars at $z \lesssim 0.3$ in an 8 hours exposure.

It should be noted, however, that the characteristics of the proposed facilities have been evolving and that some of the above predictions are therefore likely too optimistic.

5.2 UV

Cen et al. (2001) find that about 20–30 per cent of the WHIM at $z = 0$, or equivalently, about 10 per cent of the baryonic mass, is traced by O VI absorption lines with equivalent width $W > 20$ mÅ. They expect about five O VI absorption lines per unit redshift with $W > 35$ mÅ and about 0.5 per unit redshift with $W > 350$ mÅ. Fang and Bryan (2001) estimate that most O VI absorption lines are correlated to filamentary regions with overdensities of 5–100 and temperatures of a few times 10^5 K.

The equivalent width of absorption systems in the ultraviolet region can be useful to distinguish a prevalently photoionised from a prevalently collisionally ionised IGM. Collisional ionisation dominates in the denser and warmer regions of the IGM and is responsible for most lines with $W \gtrsim 40$ mÅ. On the other hand, photoionisation is predominant in the cooler and less dense regions and usually produces narrow absorption lines with $W \lesssim 40$ mÅ. Essentially all lines with equivalent width $W > 80$ mÅ are due to collisionally ionised gas (Fang and Bryan 2001).

Richter et al. (2006) investigate the origin of the broad H I Lyα absorbers (BLA) with Doppler parameters $b > 40$ km s^{-1} seen in UV spectra from the *Hubble Space Telescope* and *FUSE* (Richter et al. 2004, 2008—Chap. 4, this issue). They find that such absorbers are associated with WHIM gas with temperatures in the range $2.5 \times 10^4 < T < 1.6 \times 10^6$ K, that represents about a quarter of the gas mass in the simulation at $z = 0$. Less than one third of the broad H I absorbers are associated with cooler gas with $T < 2 \times 10^4$ K, whose broadening is produced by non-thermal processes such as turbulence and velocity structures in the IGM.

6 Sunyaev–Zel'dovich Signal of the WHIM

It has been suggested that ionised WHIM gas could possibly be detected through the thermal Sunyaev–Zel'dovich effect it produces in the CMB (e.g., Croft et al. 2006; Cao et al. 2006; Hallman et al. 2007; Bregman 2007). When the CMB radiation crosses a hot ionised medium such as the WHIM, the energy of the passing photons can be increased by collisions with the free electrons in the WHIM. This produces a local increase of the CMB temperature, with an increase proportional to the line integral of the pressure of the hot gas. Although the signal arising from the hot gas in large clusters has been clearly identified, a weaker signal associated with WHIM gas has not yet been reliably detected in current CMB data (Hernandez-Monteagudo et al. 2004; Hansen et al. 2005).

Hallman et al. (2007) estimate that the WHIM could contribute on average between 4 and 12 per cent of the integrated Sunyaev–Zel'dovich effect (SZE) signal Y for individual sources identified in upcoming surveys. The WHIM is a particularly relevant source of contamination for observations of clusters of galaxies, because it would strongly affect the calibration of the cluster Y–M relationship. Unless the WHIM signal can be unambiguously modelled and separated from the cluster signal, it introduces a bias and a source of scatter difficult to quantify in the Y–M relationship. Ultimately, this would limit the reliability of the Y–M relationship for estimating cosmological parameters.

7 Intergalactic Magnetic Fields

It has been suggested that cosmic shear and gravitational shocks might be able to amplify the magnetic fields that permeate the large-scale structure (e.g., Dolag et al. 1999; Miniati et al. 2001; Brüggen et al. 2005). Given the strong interconnection between the origin of the WHIM and gravitational shocks, it is possible for the WHIM to be magnetised, although it might be difficult to predict to what level.

While magnetic fields have been unambiguously detected in the cores of galaxy clusters, few observations have attempted to establish the presence of magnetised plasma in low density regions such as groups and filaments. However, there is now compelling evidence that magnetic fields might exist beyond clusters, and in particular in the WHIM. Kim et al. (1989) claim the detection of an extended magnetic field in the region of the Coma supercluster. They observe radio-synchrotron radiation from a region with an enhanced number density of galaxies, seeming to indicate the presence of a group of galaxies, perhaps in the process of merging with the Coma cluster. Evidence for intergalactic magnetic fields in the region ZwCL 2341+0000 at $z \sim 0.3$, where galaxies seems to lie along a filament, has been reported by Bagchi et al. (2002). In both cases, the inferred strength of the magnetic field is of order 0.01–0.1 μG.

Several mechanisms have been proposed to explain the injection of seed magnetic fields into the intracluster and intergalactic media, later to be amplified by cosmic shear and gravitational shocks. Among them, primordial magnetic fields, jets and radio lobes emerging from radio galaxies (e.g., Hoyle 1969; Chakrabarti et al. 1994; Enßlin et al. 1997) and galactic winds (e.g., Kronberg et al. 1999; Bertone et al. 2006). Simulations (Kronberg et al. 1999; Kronberg et al. 2001; Bertone et al. 2006) show that magnetic fields can be efficiently transported into the IGM and fill most of the universe by $z = 0$. Bertone et al. (2006) predict that galactic winds alone could be responsible for IGM magnetic fields of order 10^{-6}–10^{-2} μG, in agreement with observations.

8 Conclusions

The search for the "missing baryons" has opened the way to our understanding of the most elusive baryonic component in the universe: the hot diffuse gas that traces the cosmic web.

Given the paucity of observational findings, most of what we know about the WHIM is based on numerical simulations. These predict that at $z < 1$ up to 50 per cent of the IGM has been heated to temperatures in the range $T \sim 10^5$–10^7 K. While most of the heating is provided by gravitational accretion shocks propagating out of regions undergoing gravitational collapse, non-gravitational processes such as galactic winds and AGN feedback may play an important role in high density regions, mostly by preventing cooling and regulating the intensity of the X-ray background.

The observability of the WHIM by future facilities has received a lot of attention and a wealth of predictions have been made based on numerical results. The detection of the WHIM seems to be extremely challenging, if not impossible, with current instruments and a new generation of X-ray telescopes with higher sensitivity and spectral resolution, such as *Constellation-X* and *XEUS*, is needed to study the WHIM in absorption. In addition to high sensitivity and spectral resolution, high spatial resolution on a large field of view, as proposed for the Explorer of Diffuse emission and Gamma-ray burst Explosions (*EDGE*), would be ideal for mapping the WHIM emission on the sky.

Although most theoretical predictions claim optimistically that we will be able to detect the WHIM in absorption with *Constellation-X* and *XEUS*, and maybe in emission with some

luck, there is a long way to go before observations can confirm the simulation predictions, or show us a different picture altogether of what and where the WHIM is.

Acknowledgements We are grateful to Frits Paerels and Philipp Richter for a careful reading of the manuscript. We thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". This work was supported by a Marie Curie Excellence Grant MEXT-CT-2004-014112. SB acknowledges support from STFC and from a NSF grant AST-0507117.

References

J. Bagchi, T.A. Enßlin, F. Miniati et al., New Astron. **7**, 249 (2002)
X. Barcons, S. Mateos, M.T. Ceballos, Mon. Not. R. Astron. Soc. **361**, 13 (2000)
S. Bertone, C. Vogt, T.A. Enßlin, Mon. Not. R. Astron. Soc. **370**, 319 (2006)
M. Bonamente, R. Lieu, J.P.D. Mittaz, Astrophys. J. **547**, L7 (2001)
J.N. Bregman, Annu. Rev. Astron. Astrophys. **45**, 221 (2007)
G.L. Bryan, G.M. Voit, Astrophys. J. **556**, 590 (2001)
M. Brüggen, M. Ruszkowski, A. Simionescu, M. Hoeft, C. Dalla Vecchia, Astrophys. J. **631**, L21 (2005)
A.M. Bykov, K. Dolag, F. Durret, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9312-9
L. Cao, W. Zuo, Y.-Q. Chu, Mod. Phys. Lett. A **21**, 2233 (2006)
R. Cen, J.P. Ostriker, Astrophys. J. **514**, 1 (1999)
R. Cen, T.M. Tripp, J.P. Ostriker, E.B. Jenkins, Astrophys. J. **559**, L5 (2001)
R. Cen, T. Fang, Astrophys. J. **650**, 573 (2006)
R. Cen, J.P. Ostriker, Astrophys. J. **650**, 560 (2006)
S.K. Chakrabarti, R. Rosner, S.I. Vainshtein, Nature **368**, 434 (1994)
X. Chen, D.H. Weinberg, N. Katz, R. Davé, Astrophys. J. **594**, 42 (2003)
L.-M. Cheng, S. Borgani, P. Tozzi et al., Astron. Astrophys. **431**, 405 (2005)
R.A.C. Croft, A.J.R. Banday, L. Hernquist, Mon. Not. R. Astron. Soc. **369**, 1090 (2006)
R.A.C. Croft, T. Di Matteo, R. Davé et al., Astrophys. J. **557**, 67 (2001)
C.W. Danforth, J.M. Shull, Astrophys. J. (2007, submitted). astro-ph/0709.4030
R. Davé, L. Hernquist, N. Katz, D.H. Weinberg, Astrophys. J. **511**, 521 (1999)
R. Davé, R. Cen, J.P. Ostriker et al., Astrophys. J. **552**, 473 (2001)
K. Dolag, M. Meneghetti, L. Moscardini, E. Rasia, A. Bonaldi, Mon. Not. R. Astron. Soc. **370**, 656 (2006)
K. Dolag, M. Bartelmann, H. Lesch, Astron. Astrophys. **348**, 351 (1999)
K. Dolag, S. Borgani, S. Schindler, A. Diaferio, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9316-5
F. Durret, J. Kaastra, J. Nevalainen, T. Ohashi, N. Werner, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9313-8
T.A. Enßlin, P.L. Biermann, P.P. Kronberg, X. Wu, Astrophys. J. **477**, 560 (1997)
T. Fang, G.L. Bryan, Astrophys. J. **561**, L31 (2001)
T. Fang, G.L. Bryan, C.R. Canizares, Astrophys. J. **564**, 604 (2002)
T. Fang, R.A.C. Croft, W.T. Sanders et al., Astrophys. J. **623**, 612 (2005)
A. Finoguenov, G.U. Briel, P.J. Henry, Astron. Astrophys. **410**, 777 (2003)
M. Fukugita, P.J.E. Peebles, Astrophys. J. **616**, 643 (2004)
M. Fukugita, C.J. Hogan, P.J.E. Peebles, Astrophys. J. **503**, 518 (1998)
S.R. Furlanetto, A. Loeb, Astrophys. J. **611**, 642 (2004)
S.R. Furlanetto, J. Schaye, V. Springel, L. Hernquist, Astrophys. J. **606**, 221 (2004)
S.R. Furlanetto, L.A. Phillips, M. Kamionkowski, Mon. Not. R. Astron. Soc. **359**, 295 (2005)
E.J. Hallman, B.W. O'Shea, J.O. Burns et al., Astrophys. J. **671**, 27 (2007)
F.K. Hansen, E. Branchini, P. Mazzotta, P. Cabella, K. Dolag, Mon. Not. R. Astron. Soc. **361**, 753 (2005)
G. Hasinger, R. Burg, R. Giacconi et al., Astron. Astrophys. **329**, 482 (1998)
U. Hellsten, N.Y. Gnedin, J. Miralda-Escudé, Astrophys. J. **509**, 56 (1998)
C. Hernandez-Monteagudo, R. Genova-Santos, F. Atrio-Barandela, Astrophys. J. **613**, L89 (2004)
F. Hoyle, Q.J.R. Astron. Soc. **10**, 10 (1969)
H. Kang, D. Ryu, R. Cen, D. Song, Astrophys. J. **620**, 21 (2005)
H. Kawahara, K. Yoshikawa, S. Sasaki et al., Publ. Astron. Soc. Jpn. **58**, 657 (2006)
K.-T. Kim, P.P. Kronberg, G. Giovannini, T. Venturi, Nature **341**, 720 (1989)
A.V. Kravtsov, A. Klypin, Y. Hoffman, Astrophys. J. **571**, 563 (2002)
P.P. Kronberg, H. Lesch, U. Hopp, Astrophys. J. **511**, 56 (1999)
P.P. Kronberg, Q.W. Dufton, H. Li, S.A. Colgate, Astrophys. J. **560**, 178 (2001)

R. Lieu, J.P.D. Mittaz, S. Bowyer et al., Astrophys. J. **458**, L5 (1996)
F. Miniati, T.W. Jones, H. Kang, D. Ryu, Astrophys. J. **562**, 233 (2001)
J.P.D. Mittaz, R. Lieu, R. Cen, M. Bonamente, Astrophys. J. **617**, 860 (2004)
T. Miyaji, G. Hasinger, M. Schmidt, Astron. Astrophys. **353**, 25 (2000)
B.B. Nath, J. Silk, Mon. Not. R. Astron. Soc. **327**, L5 (2001)
K.A. Olive, G. Steigman, T.P. Walker, Phys. Rev. **333**, 389 (2000)
F.B.S. Paerels, J.S. Kaastra, T. Ohashi et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9323-6
U.L. Pen, Astrophys. J. **510**, L1 (1999)
R. Perna, A. Loeb, Astrophys. J. **503**, L135 (1998)
M. Persic, P. Salucci, Mon. Not. R. Astron. Soc. **258**, 14 (1992)
L.A. Phillips, J.P. Ostriker, R. Cen, Astrophys. J. **554**, L9 (2001)
M. Pierre, G. Bryan, R. Gastaud, Astron. Astrophys. **356**, 403 (2000)
C.E. Rakowski, Adv. Space Res. **35**, 1017 (2005)
Y. Rasera, R. Teyssier, Astron. Astrophys. **445**, 1 (2006)
M. Rauch, J. Miralda-Escudé, W.L.W. Sargent et al., Astrophys. J. **489**, 7 (1997)
P. Richter, B.D. Savage, T.M. Tripp, K.R. Sembach, Astrophys. J. Suppl. Ser. **153**, 165 (2004)
P. Richter, T. Fang, G.L. Bryan, Astron. Astrophys. **451**, 767 (2006)
P. Richter, F.B.S. Paerels, J.S. Kaastra, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9325-4
M. Roncarelli, L. Moscardini, P. Tozzi et al., Mon. Not. R. Astron. Soc. **368**, 74 (2006)
J. Schaye, Astrophys. J. **559**, 507 (2001)
D.N. Spergel, R. Bean, O. Doré et al., Astrophys. J. Suppl. Ser. **170**, 377 (2007)
R. Sunyaev, Y. Zel'dovich, Astron. Astrophys. **20**, 189 (1972)
R. Teyssier, J.-P. Chièze, J.-M. Alimi, Astrophys. J. **509**, 62 (1998)
E. Ursino, M. Galeazzi, Astrophys. J. **652**, 1085 (2006)
P. Valageas, R. Schaeffer, J. Silk, Astron. Astrophys. **388**, 741 (2002)
M. Viel, E. Branchini, R. Cen et al., Mon. Not. R. Astron. Soc. **341**, 792 (2003)
M. Viel, E. Branchini, R. Cen et al., Mon. Not. R. Astron. Soc. **360**, 1110 (2005)
G.M. Voit, A.E. Evrard, G.L. Bryan, Astrophys. J. **548**, L123 (2001)
G.M. Voit, G.L. Bryan, Astrophys. J. **551**, L139 (2001)
D.H. Weinberg, J. Miralda-Escude, L. Hernquist, N. Katz, Astrophys. J. **490**, 564 (1997)
N. Yoshida, S.R. Furlanetto, L. Hernquist, Astrophys. J. **618**, L91 (2005)
K. Yoshikawa, N.Y. Yamasaki, Y. Suto et al., Publ. Astron. Soc. Jpn. **55**, 879 (2003)
K. Yoshikawa, K. Dolag, Y. Suto et al., Publ. Astron. Soc. Jpn. **56**, 939 (2004)
K. Yoshikawa, S. Sasaki, Publ. Astron. Soc. Jpn. **58**, 641 (2006)
M.A. Worsley, A.C. Fabian, F.E. Bauer et al., Mon. Not. R. Astron. Soc. **357**, 1281 (2005)
K.K.S. Wu, A.C. Fabian, P.E.J. Nulsen, Mon. Not. R. Astron. Soc. **324**, 95 (2001)

Chapter 15
Non-Thermal Processes in Cosmological Simulations

K. Dolag · A.M. Bykov · A. Diaferio

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9319-2

Abstract Non-thermal components are key ingredients for understanding clusters of galaxies. In the hierarchical model of structure formation, shocks and large-scale turbulence are unavoidable in the cluster formation processes. Understanding the amplification and evolution of the magnetic field in galaxy clusters is necessary for modelling both the heat transport and the dissipative processes in the hot intra-cluster plasma. The acceleration, transport and interactions of non-thermal energetic particles are essential for modelling the observed emissions. Therefore, the inclusion of the non-thermal components will be mandatory for simulating accurately the global dynamical processes in clusters. In this review, we summarise the results obtained with the simulations of the formation of galaxy clusters which address the issues of shocks, magnetic field, cosmic ray particles and turbulence.

Keywords Cosmology: theory, large-scale structure of universe · Acceleration of particles · Hydrodynamics · Magnetic fields · Method: numerical, N-body simulations

1 Introduction

Within the intra-cluster medium (ICM), there are several processes which involve non-thermal components, such as the magnetic field and cosmic rays (CRs). At first glance,

K. Dolag (✉)
Max-Planck-Institut für Astrophysik, P.O. Box 1317, 85741 Garching, Germany
e-mail: kdolag@mpa-garching.mpg.de

A.M. Bykov
A.F. Ioffe Institute of Physics and Technology, St. Petersburg 194021, Russia
e-mail: byk@astro.ioffe.ru

A. Diaferio
Dipartimento di Fisica Generale "Amedeo Avogadro", Università degli Studi di Torino, Torino, Italy

A. Diaferio
Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy

many of these processes can be studied numerically in simplified configurations where one can learn how the individual processes work in detail. For example, there are investigations which consider magneto-hydro-dynamical (MHD) simulations of cloud-wind interactions (see Gregori et al. 2000, and references therein) or simulate the rise of relic radio bubbles (see Jones and De Young 2005; Reynolds et al. 2005, and references herein). Such work usually focuses on the relevance of CRs, turbulence and the local magnetic fields within and around these bubbles. In this review, we will instead concentrate on simulations of non-thermal phenomena, which aim at understanding the relevance of these phenomena for galaxy cluster properties or at unveiling possible origins of the non-thermal radiation. So far, firm evidence for the presence of non-thermal emission (at radio wavelengths) and for the presence of cosmologically relevant extended magnetic fields has been found only in galaxy clusters; in filaments only weak limits have been found so far. For a more detailed discussion see the reviews on radio observations by Rephaeli et al. (2008)—Chap. 5, this issue, and Ferrari et al. 2008—Chap. 6, this issue.

2 Possible Origins of Non-thermal Components within the Large-Scale Structure (LSS)

The origin of magnetic fields and of CRs in galaxy clusters is still under debate. The variety of possible contributors to magnetic fields ranges from primordial fields, battery and dynamo fields to all the classes of astrophysical objects which can contribute with their ejecta. The latter possibility is supported by the observation of metal enrichment of the ICM, which is due to the ejecta of stars in galaxies, see Werner et al. 2008—Chap. 16, this issue, and Schindler and Diaferio 2008—Chap. 17, this issue. The magnetic fields produced by all these contributors will be compressed and amplified by the process of structure formation. The exact amount of amplification and the resulting filling factor of the magnetic field will depend on where and when the contributor is thought to be more efficient.

CRs are produced at shocks, both in supernova remnants and in cosmological accretion and merger shocks. Furthermore, active galactic nuclei (AGN) and radio galaxies contribute to the CR population. CR protons can also produce CR electrons via hadronic reactions with the thermal plasma, leading to the so-called secondary CRs.

2.1 Possible Origins of Magnetic Fields

Magnetic fields can be produced either at relatively low redshift ($z \sim 2-3$) or at high redshift ($z \gtrsim 4$). In the former case, galactic winds (e.g. Völk and Atoyan 2000) or AGN ejecta (e.g. Enßlin et al. 1997; Furlanetto and Loeb 2001, and references therein) produce magnetic fields 'locally', e.g. within the proto-cluster region. In the latter case, the magnetic field seeds can also be produced by an early population of dwarf starburst galaxies or AGN before galaxy clusters form gravitationally bound systems.

One of the main arguments in favour of the 'low-redshift' models is that the high metallicity observed in the ICM suggests that a significant enrichment occurred in the past due to galactic winds or AGN. Winds and jets should carry magnetic fields together with the processed matter. It has been shown that winds from ordinary galaxies give rise to magnetic fields which are far weaker than those observed in galaxy clusters, whereas magnetic fields produced by the ejecta of starburst galaxies can be as large as 0.1 μG. Clearly, this class of models predicts that magnetic fields are mainly concentrated in and around galaxies and within galaxy clusters. Note that if the magnetic pollution happens early enough

(around $z \sim 3$), these fields will be amplified not only by the adiabatic compression of the proto-cluster region, but also by shear flows, turbulent motions, and merging events during the formation of the cluster. Shocks are expected to be produced copiously during the non-linear stage of the LSS formation (see below). Recent detailed studies of shock propagation revealed the presence of specific instabilities driven by energetic accelerated particles (Bell and Lucek 2001; Vladimirov et al. 2006). Such instabilities result in a strong, non-adiabatic amplification of an upstream magnetic field seed which converts an appreciable fraction of the shock ram pressure into magnetic fields.

In the 'high-redshift' models of the magnetic field generation, the strength of the field seed is expected to be considerably smaller than in the previous scenario, but the adiabatic compression of the gas and the shear flows driven by the accretion of structures can give rise to a considerable amplification of the magnetic fields. Several mechanisms have been proposed to explain the origin of magnetic field seeds at high redshift. Some of them are similar to those discussed above, differing only in the time when the magnetic pollution is assumed to take place. In the present class of models the magnetic field seeds are supposed to be expelled by an early population of AGN or dwarf starburst galaxies at a redshift between 4 and 6 (Kronberg et al. 1999); this process would magnetise a large fraction of the volume. Recently, the validity of such a scenario has been confirmed by a semi-analytic modelling of galactic winds (Bertone et al. 2006).

Alternative models invoke processes that took place in the early universe. Indeed, the presence of magnetic fields in almost all the regions of the universe suggests that they may have a cosmological origin. In general, most of the 'high-z models' predict magnetic field seeds filling the entire volume of the universe. However, the assumed coherence length of the field crucially depends on the details of the model. While scenarios based on phase transitions give rise to coherence lengths which are so small that the corresponding fields have probably been dissipated, magnetic fields generated at neutrino or photon decoupling are thought to have a much higher chance of surviving until the present time. Another (speculative) possibility is that the field seed was produced during inflation. In this case, the coherence length can be as large as the Hubble radius. See Grasso and Rubinstein (2001) for a review.

Magnetic field seeds can also be produced by the so-called Biermann battery effect (Kulsrud et al. 1997; Ryu et al. 1998). The idea here is that merger/accretion shocks related to the hierarchical structure formation process give rise to small thermionic electric currents which, in turn, may generate magnetic fields. The battery process has the attractive feature of being independent of unknown physics at high redshift. Its drawback is that, due to the large conductivity of the ICM, it can give rise to very tiny magnetic fields, of order 10^{-21} G at most. One therefore needs to invoke a subsequent turbulent dynamo to boost the field strength to the observed level.

Such a turbulent amplification, however, cannot be simulated numerically yet, making it quite difficult to predict how it would proceed in a realistic environment. It is clear that one expects the level of turbulence to be strongly dependent on the environment, and that it should appear mostly in high-density regions like collapsed objects. While energetic events, such as mergers of galaxy clusters, can easily be effective at driving the required levels of turbulence, it is harder to understand how turbulent amplification would work in relatively quiet regions like filaments. Here, magnetic fields can be rather amplified by the shocks which originate during the LSS formation, as mentioned at the beginning of this section. Lacking a theoretical understanding of the turbulent amplification, it is therefore not straightforward to relate the very weak field seeds produced by the battery process to the

magnetic fields observed today. Attempts to construct such models, based on combining numerical and analytical computations, have not yet been reported to successfully reproduce the observed scaling relations of the magnetic fields in galaxy clusters.

2.2 Possible Origins of CRs

There are many possible contributors to the CR population in the ICM. AGN and stellar activity in cluster galaxies are believed to be a significant source of CRs in the ICM. Furthermore, cluster accretion shocks of high Mach numbers (typically above 100) and of sizes of megaparsec or larger are widely accepted to be possible CR sources with nuclei accelerated up to $\sim 10^9$ GeV (e.g. Norman et al. 1995). Cluster mergers generate internal shocks (of moderate Mach numbers, typically below 4) which provide most of the ICM gas heating (see Kang et al. 2005), but are also likely to convert a non-negligible fraction ($\gtrsim 10$ %) of their power into CRs.

AGN outflows dissipate their kinetic energy and Poynting fluxes into the ICM providing additional non-gravitational ICM heating and, thus, a plausible feedback solution to the cooling flow problem (e.g. Binney and Tabor 1995; Churazov et al. 2003). Recent *XMM-Newton* observations of the cluster MS 0735+7421 (Gitti et al. 2007) revealed that a powerful AGN outburst deposited about 6×10^{61} erg into the ICM outside the cooling region of radius ~ 100 kpc. Relativistic outflows of AGN are likely to be an essential source of superthermal particles in clusters. A powerful relativistic AGN jet could deposit up to 10^{62} erg into a relativistic particle pool during a duty cycle of about ~ 50 Myr. A potential source of energetic particles are sub-relativistic ions with energy ~ 100 MeV that are evaporated from an AGN, which accretes mass in the ion-supported tori regime (e.g. Rees et al. 1982). A certain fraction of the energetic particles could escape the flows and avoid fast cooling due to collective effects in the central parts of the cluster. Re-acceleration of that non-thermal population by inner shocks inside the cluster can provide a long-lived non-thermal component that contributes to the total pressure of the ICM.

Star formation activity in galaxies is another source of CRs in clusters of galaxies (e.g. Völk et al. 1996, and references therein). The combined action of supernovae and winds of early-type stars leads to the formation of a hot, X-ray emitting, slowly expanding superbubble filled with large-scale (tens of parsecs) compressible MHD motions (e.g. rarefactions) and shocks. When the starburst event is energetic enough, the superbubble may expand beyond the disk of the parent galaxy and produce a superwind that supplies the intergalactic medium with metals ejected by supernovae and CRs. The superwinds have been seen with Chandra in some starburst galaxies: M 82, Arp 220 and NGC 253. In the superwind scenario, the CR ejection is tightly connected with the metal enrichment of the cluster (Schindler and Diaferio 2008—Chap. 17, this issue).

Bykov (2001) argues that non-thermal particle acceleration can favourably occur in correlated supernovae events and powerful stellar winds with great energy release, which generate interacting shock waves within superbubbles. The acceleration mechanism provides efficient creation of a non-thermal nuclei population with a hard low-energy spectrum; this population can transport a substantial fraction ($\sim 30\%$) of the kinetic energy released by the supernovae and by the winds of young massive stars.

A bright phase in the galaxy evolution can be the source of the relic CRs in clusters. Energetic nuclei can be stored in cluster magnetic fields for several Hubble times (e.g. Berezinsky et al. 1997). The presence of these nuclei produces a diffuse flux of high-energy γ and neutrino radiation due to the interaction of the CRs with the ICM (Völk et al. 1996; Enßlin et al. 1997). The resulting flux depends on the ICM baryon density.

A physical model of particle acceleration by the ensemble of inner shocks in the ICM could be similar to the superbubble model. The sub-cluster merging processes and the supersonic motions of dark matter halos in the ICM are accompanied by the formation of shocks, large-scale flows and broad spectra of MHD-fluctuations in a tenuous intra-cluster plasma with frozen-in magnetic fields. Vortex electric fields generated by the large-scale motions of highly conductive plasma harbouring shocks result in a non-thermal distribution of the charged nuclei. The free energy available for the acceleration of energetic particles is in the ram pressure of the shocks and in the large-scale motions.

The most studied way to transfer the power of the MHD motions to the energetic particle population is the Fermi-type acceleration (see e.g. the review by Blandford and Eichler 1987 and Bykov et al. 2008—Chap. 7, this issue). An important ingredient of the energetic particle acceleration by shocks and large-scale MHD motions is the presence of small-scale MHD turbulence, which is necessary to scatter relativistic particles and to make their pitch-angle isotropic. The scale of the fluctuations required for the resonant scattering of a particle of energy ~ 1 GeV is about $3\times 10^{12} B_{-6}^{-1}$ cm, where B_{-6} is the local mean magnetic field in μG. The scale is some 10–11 orders of magnitude smaller than the basic energy scale of the system; thus the origin and the maintenance mechanisms of such small scale turbulence is a serious problem.

In non-linear models of particle acceleration by strong MHD shocks the presence of turbulence could be supported by the CR instabilities themselves (see e.g. recent non-linear Monte-Carlo simulations by Vladimirov et al. 2006). Direct MHD cascade of energy to small scale turbulence is also possible, but the small scale fluctuations are highly anisotropic in that case (see e.g. Biskamp 2003). Moreover, the cascade properties are still poorly known at the scales close to the Coulomb mean free path where the transition from the collisional to the collisionless plasma regime occurs.

The particle distribution within a system with multiple shocks and with large-scale plasma motions is highly intermittent, because it has strong peaks at shocks and a non-steady mean CR background. By using the kinetic equations given by Bykov (2001), one can construct a model of the temporal evolution of the particle distribution function, which accounts for the non-linear effect of the reaction of the accelerated particles on the shock turbulence inside the cluster. In Fig. 1 we show the efficiency of a conversion into CRs of the power of shocks and the bulk plasma motions of a scale of about 100 kpc. One may note that, while the efficiency could be as high as 30 % for some relatively short period of time, the efficiency is typically ~ 10 % for most of the time. Fig. 2 shows the temporal behaviour of the particle spectra on a Gyr time scale.

Non-thermal particles and magnetic fields could contribute to the total pressure of the ICM. In the particular model described above a substantial energy density could be stored in protons of energy below 200 MeV. Protons of that energy level produce very little high energy emission and therefore it is rather difficult to constrain that component directly from the observational data which are currently available.

3 Shocks in the LSS and Clusters

As described in the previous section, shocks can play a crucial role as source of non-thermal components in galaxy clusters. Therefore, a significant effort is spent to investigate the distribution and the evolution of shocks during structure formation.

In early simulations of the formation of cosmological structures, it was already noticed that shocks are the main source of the ICM heating (see Quilis et al. 1998, and references therein). In numerical simulations, two classes of shocks can be identified.

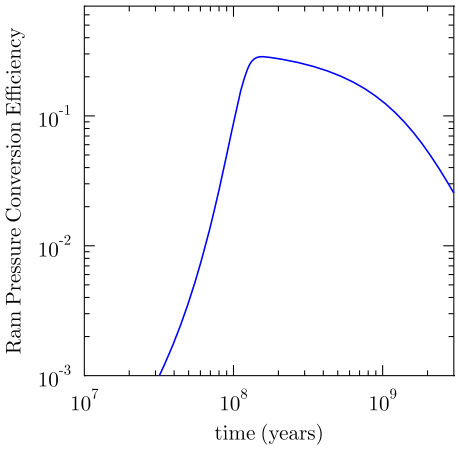


Fig. 1 The simulated efficiency of conversion of shock and bulk MHD motions ram pressure power to the CRs at different stages of cluster evolution. The efficiency is defined as a volume averaged fraction of the cluster large-scale motions power transferred to CRs. The simulation was done for a model of CR acceleration by an ensemble of shocks and MHD-turbulent motions described in Bykov (2001), adopted for clusters of galaxies. The MHD turbulence in a cluster of a Mpc size was initiated with the Gaussian bulk velocity dispersion of 2 000 km s^{-1} at a scale of 100 kpc

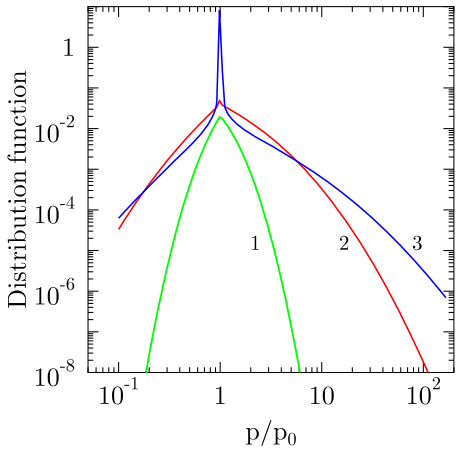


Fig. 2 The temporal evolution of the particle distribution function in a cluster of a Mpc size and a Gaussian bulk velocity dispersion of about 2 000 km s^{-1} at a scale of 100 kpc. The curves labelled as *1, 2, 3* correspond to 0.1, 0.3 and 1 Gyr age of the cluster. Mono-energetic proton injection at energy 10 keV (p_0 is the injection momentum). The injection efficiency (see Fig. 1) was assumed to be about 10^{-3}. The power conversion efficiency for the model was already illustrated in Fig. 1

The first one is the accretion shocks surrounding the forming objects, namely filaments or galaxy clusters. In the case of galaxy clusters, most shocks are quasi-spherical and their position with respect to the centre of the cluster grows with time (shock travelling outwards). These shocks can be associated with the accretion of diffuse matter which is relatively cold. These shocks provide a first step towards the virialisation of an initially supersonic accreting flow. The radius of the accretion shock (for a quasi-spherical cluster at $z = 0$) is typically two times the virial radius. As the typical temperature and density of the upstream gas is quite low, such shocks can have very large Mach numbers ($\gg 10$).

The second class of shocks in galaxy clusters arises from the merging of substructures. The gas in gravitationally bound substructures is much denser than the diffuse ICM; there-

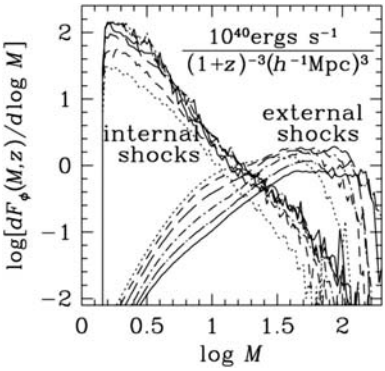


Fig. 3 Shocks detected in a cosmological simulation of a cubic volume with a side length of 100 Mpc, resolved with 512^3 grid points. The figure shows the kinetic energy flux per unit comoving volume through surfaces of external and internal shocks with Mach number between $\log M$ and $\log M + \mathrm{d}(\log M)$ at different redshifts (*solid*, $z = 0$ to *dotted*, $z = 2$). From Ryu et al. (2003)

fore it cannot be stopped in the accretion shock and it continues to move with the substructure through the gas, which is mostly virialised. The substructures, which have velocities larger than the sound velocity of the hot ICM, drive the so-called merger shocks. These shocks are initiated within the ICM which has been already shock-heated, and therefore they typically have modest Mach numbers (below $\approx$4). The merger shocks propagate within the dense ICM and are likely to be the main contributor to the global energy dissipation within the ICM (see Pfrommer et al. 2006, and references therein). As these shocks are primarily driven by gravitational attraction, non-gravitational processes (like cooling or feedback by star formation processes) do not significantly affect the global energy dissipation (see Kang et al. 2007). The ram pressure dissipation process in shocks within the collisionless plasma is non-trivial because of the role of long-lived non-thermal components, namely magnetic fields, MHD-waves and energetic particles (see for details Bykov et al. 2008—Chap. 7, this issue). The complex multi-scale structure of a shock modified by energetic CRs (e.g. where the back reaction of the CRs on the thermal plasma is strong enough to change the thermodynamical states of the plasma in addition to the shock itself) cannot be resolved properly in LSS simulations.

The efficiency of energising CRs by shocks strongly depends on a number of factors. Among the most important factors, we have (i) the upstream magnetic field of the shock, (ii) the plasma parameter $\beta = \mathcal{M}_\mathrm{a}^2/\mathcal{M}_\mathrm{s}^2$ (ratio of thermal to magnetic pressure, here expressed by the Mach numbers $\mathcal{M}_\mathrm{a}$ and $\mathcal{M}_\mathrm{s}$ with respect to the Alfvén and sound velocity respectively), (iii) the inclination of the upstream field of the shock, and (iv) the presence of a pre-existing CR population produced by the accretion shock or in star-forming regions.

The distribution of the Mach numbers of the shocks within the cosmological structure can be derived from semi-analytical modelling (see Gabici and Blasi 2003; Pavlidou and Fields 2006, and references therein), as well as directly from cosmological simulations (see Miniati et al. 2000; Ryu et al. 2003; Pfrommer et al. 2006; Kang et al. 2007, and references therein). An example is shown in Fig. 3. There are significant differences in the strength and the distribution of the shocks found in different works: this fact has strong consequences on the predicted amount and the relative importance of CRs and their energy with respect to the thermal energy of the ICM. Note that part of the confusion, specially in the early work, is caused by a lack of resolution within the simulations, which prevents internal shocks from being properly resolved (see Ryu et al. 2003; Pfrommer et al. 2006); further confusion originates from the influence of other, non-thermal processes (like preheating) on the inferred Mach number of the external shocks (see Pfrommer et al. 2006; Kang et al. 2007). Moreover, differences in the underlying hydrodynamic simulations as well as in the shock detection, mostly done in a post processing fashion, can contribute to

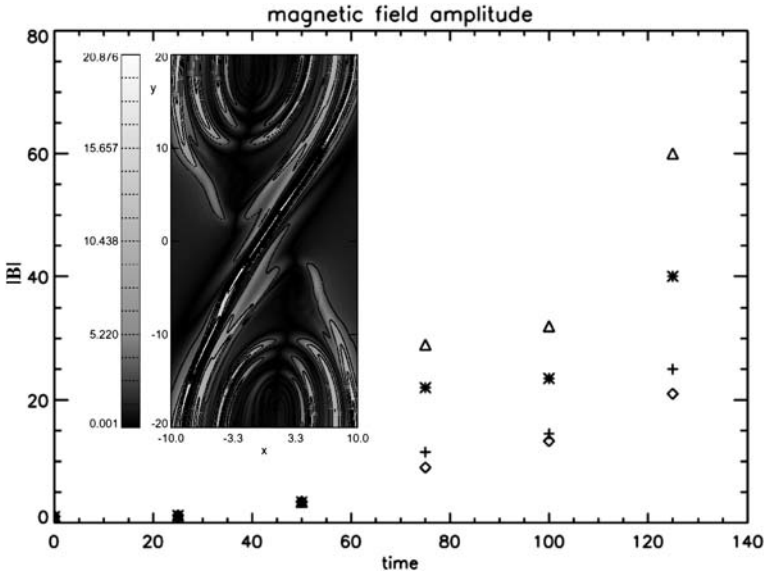


Fig. 4 Resistive, MHD simulation of a flow which is KH instable at the interface separating the left and right halve of the simulated volume. Shown is the magnetic field (y-component) in units of the initial magnetic field after evolving the KH instability (inlay) and the time evolution of the (relative) magnetic field amplitude for different values of the chosen resistivity and initial magnetic field. Time is in units of the Alfvénic transit time (4000 year). *Triangles* and *stars* are for an initial magnetic field of 10 μG, *crosses* and *diamonds* for a ten times larger field. *Stars* and *diamonds* have a magnetic Reynolds number of 100, *triangles* and *crosses* have a ten times larger magnetic Reynolds number. From Birk et al. (1999)

differences in the results. Nevertheless, there is a qualitative agreement in the case of simulated internal shocks between the simulations done by Ryu et al. (2003) and Pfrommer et al. (2006).

4 Evolution of Magnetic Fields in Simulations

4.1 Local Amplification of Magnetic Fields

A very basic process which amplifies magnetic fields is related to the Kelvin-Helmholtz (KH) instabilities driven by shear flows, which are common during the formation of cosmic structures. Birk et al. (1999) performed a detailed study of such an amplification within the outflows of starburst galaxies, where the KH timescale should be $\approx 4 \times 10^5$ years (see Fig. 4). By using a Cartesian resistive MHD code they found that the amplification factor of the magnetic field mainly depends on the initial ratio of the magnetic to the kinetic energy and only mildly depends on the assumed resistivity. They concluded that such a process could indeed explain why the magnetic field observed in the halo of starburst galaxies is significantly higher than what is expected from the magnetic fields observed within galactic disks. When applied to a cluster core environment, the KH timescale turns out to be 10^7 years; this timescale makes the KH instability an interesting process to amplify weak magnetic fields.

It is quite expected that merger events and accretion of material onto galaxy clusters will drive significant shear flows within the ICM. Extensive MHD simulations of single

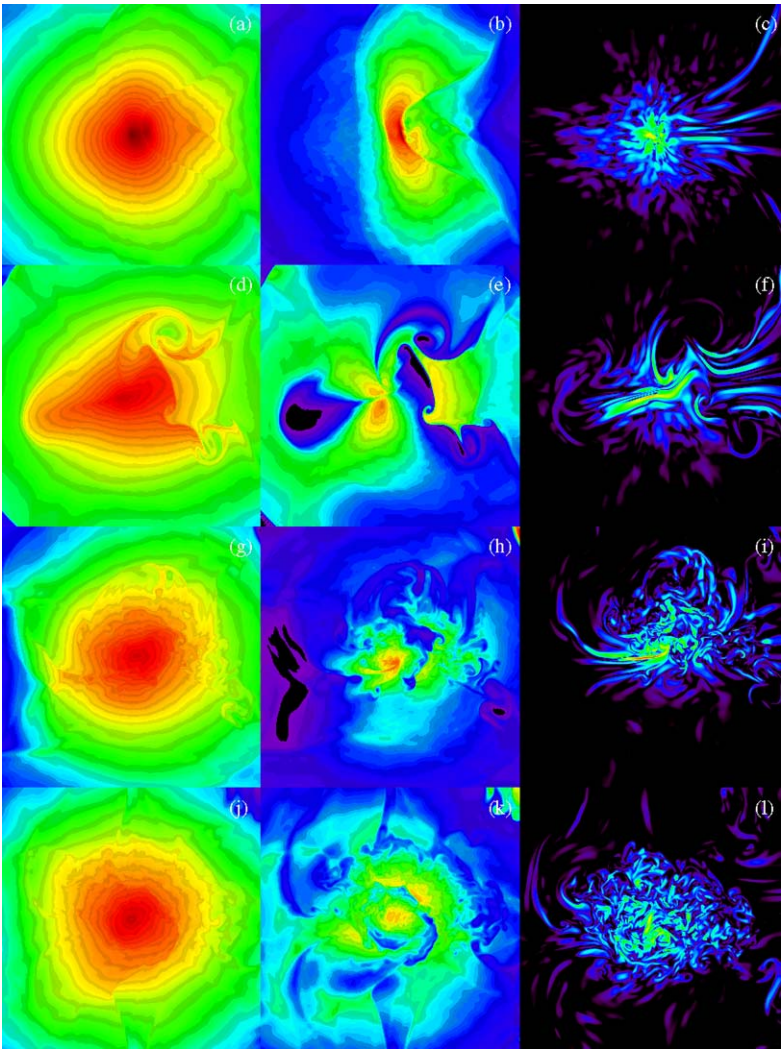


Fig. 5 The evolution of (the logarithm of) the gas density (*left column*), the gas temperature (*central column*), and (the logarithm of) the magnetic pressure (*right column*) in two-dimensional slices taken through the core of a cluster in a major merger phase in the plane of the merger. Each row refers to different epochs: $t = 0$ (i.e. the time of the core coincidence), $t = 1.3$, $t = 3.4$, and $t = 5.0$ Gyr, from *top* to *bottom*. Each panel is 3.75×3.75 Mpc. The merging subcluster enters from the *right*. From Roettiger et al. (1999b)

merging events performed using the Eulerian code *ZEUS* (Stone and Norman 1992a, 1992b) demonstrated that merger events are effective at amplifying magnetic fields (Roettiger et al. 1999b). In particular these authors found that the field initially becomes quite filamentary, as a result of stretching and compression caused by shocks and bulk flows during infall; at this stage only a minimal amplification occurs. When the bulk flow is replaced by turbulent motions (e.g., eddies), the field amplification is more rapid, particularly in localised regions, see Fig. 5. The total magnetic field energy is found to increase by nearly a factor of three

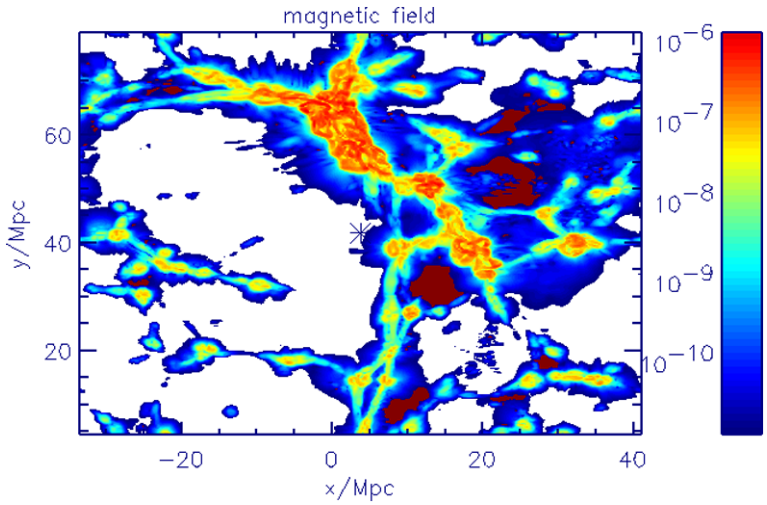


Fig. 6 A two-dimensional cut through a cosmological box simulated by Miniati et al. (2001), following the evolution of battery fields. The logarithm of the up-scaled magnetic field strength in Gauss at $z = 0$ is shown. Taken from Sigl et al. (2004)

with respect to a non-merging cluster. In localised regions (associated with high vorticity), the magnetic energy can increase by a factor of 20 or more.

A power spectrum analysis of the magnetic energy showed that the amplification is largely confined to scales comparable to or smaller than the cluster cores: this indicates that the core dimensions define the injection scale. It is worth taking notice that the previous results can be considered as a lower limit on the total amplification, because of the lack of resolution. Furthermore, it is quite likely that a galaxy cluster undergoes more than one of these events during its formation process, and that the accretion of smaller haloes also injects turbulent motions into the ICM: consequently, the magnetic field amplification within galaxy clusters will be even larger. A detailed discussion of the amplification of magnetic field in the cluster environment, using various simulations of driven turbulence, can be found in Subramanian et al. (2006); these authors show that turbulent processes can provide reasonable strength and length scales of the magnetic fields in galaxy clusters.

4.2 Magnetic Fields From Cosmological Shocks

Cosmological shocks, mainly the accretion shocks on cosmological objects like galaxy clusters and filaments, occur more frequently than the ones produced by individual merger events. They also can produce magnetic fields through the so-called Biermann battery effect (Kulsrud et al. 1997; Ryu et al. 1998), on which a subsequent turbulent dynamo may operate (see Sect. 2).

In such a scenario, the magnetic field is strongly correlated with the large-scale structure (see Fig. 6). This means that the magnetic field within the filamentary structure could be even slightly higher than its equipartition value without violating the (weak) upper limits of the rotation measure (RM) of quasars, as pointed out by Ryu et al. (1998). It is worth pointing out that the arguments for the turbulent dynamo action, which could amplify the battery field seeds up to the μG level, as presented in Kulsrud et al. (1997), refer explicitly to regions which are close to collapse into galaxies. It has still to be proven that such arguments hold

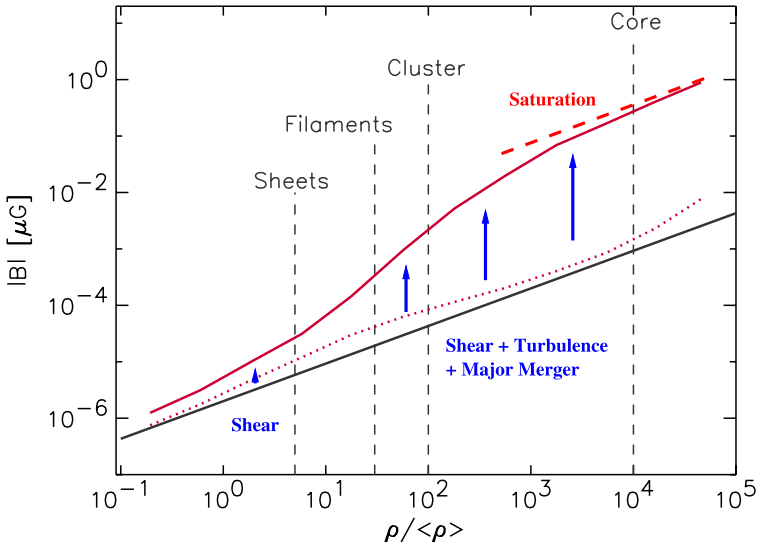


Fig. 7 The strength of the magnetic field as a function of baryonic overdensity within a cosmological simulation. The *straight line* shows the expectation for a purely adiabatic evolution, the *solid line* gives the mean field strength at a given overdensity within a cosmological simulation (Dolag et al. 2005a). While the latter is close to the adiabatic value in underdense regions, in clusters there is a significant inductive amplification due to shear flows and turbulence; this amplification however saturates in the cluster cores. At any given density, a large fraction of particles remains close to the adiabatic expectation, as shown by the *dotted line*, which gives the median of the distribution at each density

within proto-clusters or even cosmological structures, like sheets and filaments. In general, the time evolution of the magnetic field, as predicted by these simulations, strongly flattens around $z \approx 3$ (Kulsrud et al. 1997); the evolution then leads to a magnetic field intensity which is relatively uniform on scales of tens of Mpc within the LSS around galaxy clusters (see Fig. 6). Note that, so far, the synthetic rotation measures due to the magnetic fields predicted by up-scaling the battery fields have never been compared with the rotation measure observed on the scale of galaxy clusters. This might be partially motivated by the lack of resolution in the mock clusters within such non-adaptive, Eulerian simulations.

4.3 Cosmological MHD Simulations

By using the *GrapeMSPH* code (Dolag et al. 1999) and assuming that a small initial magnetic field seed exists before structure formation (see Sect. 2), the first self-consistent simulations which follow the magnetic field amplification during the formation of galaxy clusters within a cosmological environment have been performed by Dolag et al. (1999, 2002). These simulations demonstrated that the contribution to the amplification of magnetic fields by shear flows (and by its induced turbulence) is significant (see Fig. 7). Therefore, for the first time, a consistent picture of the magnetic field in galaxy clusters could be constructed: the amplification predicted by the simulations was capable of linking the predicted strength of the magnetic field seed (see Sect. 2 and references therein) at high redshift ($z \approx 3$ and higher) to the observed magnetic field strength in galaxy clusters in the local universe.

Furthermore, the simulations predicted that the final structure of the magnetic field in galaxy clusters reflects the process of structure formation, and no memory on the initial

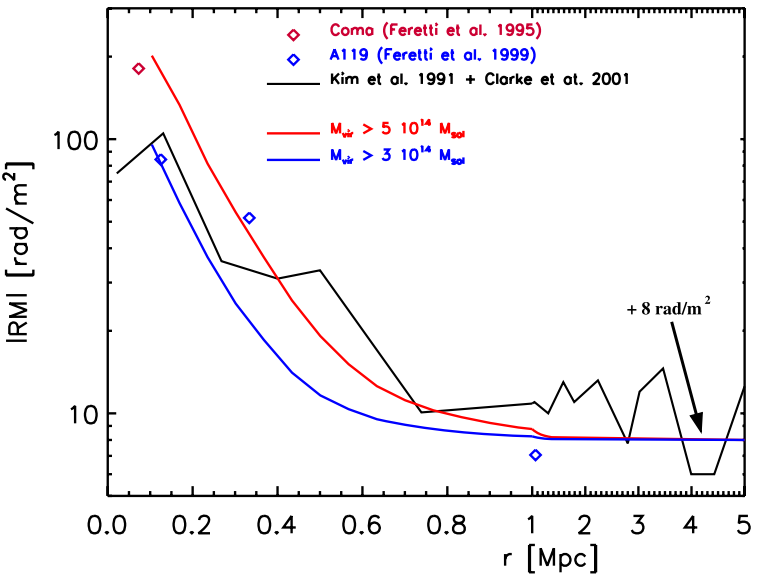


Fig. 8 Comparison of RMs in simulations with observations of Abell clusters, as a function of the distance from the cluster centre. The *smooth lines* represent the median values of |RM| produced by simulated clusters with masses above 5×10^{14} $M_\odot$ (*upper line*) and 3×10^{14} $M_\odot$ (*lower line*). The *broken line* represents the median of combined data taken from the independent samples presented in Kim et al. (1991) and Clarke et al. (2001). We also include data (*diamonds*) for the three elongated sources observed in A119 (Feretti et al. 1999), and for the elongated source observed in the Coma cluster (Feretti et al. 1995)

magnetic field configuration survives: this relaxes the constraints on models for magnetic field seeds. In general, such models predict a magnetic field profile similar to the density profile. Thereby the predicted rotation measure (RM) profile agrees with the observed one (see Fig. 8). Dolag et al. (2001) found a quasi linear correlation between two observables, namely the X-ray surface brightness and the RM r.m.s., both in observations and in their MHD simulations. By extending the *GADGET2* code to follow the full set of ideal MHD equations, Dolag et al. (2004, 2005a) performed several realisations of a cosmological volume which confirm these results at even much higher resolution.

Recently, Brüggen et al. (2005) confirmed all the previous results, which are based on SPH codes, with a simulation of the formation of a single galaxy cluster in a cosmological framework, using a passive MHD solver implemented into FLASH, an adaptive mesh refinement (AMR) code (see Fig. 9).

Another interesting quantity to look at is the slope α of the magnetic field power spectrum ($\propto k^{-\alpha}$, with k being the wave vector). Within galaxy clusters, α is predicted by the SPH simulations (Dolag et al. 2002; Rordorf et al. 2004) to be slightly lower, but still very close to $11/3$, which is the expected value for a Kolmogorov-like spectrum in 3D. The AMR simulation by Brüggen et al. (2005) nearly perfectly matches the Kolmogorov slope.

4.4 Approximative Simulations

The fully consistent cosmological MHD simulations described above are very expensive in their requirements of computing resources; thus, for the investigation of various models, simplified approaches are often used in astrophysics. Recently, Bertone et al. (2006) used

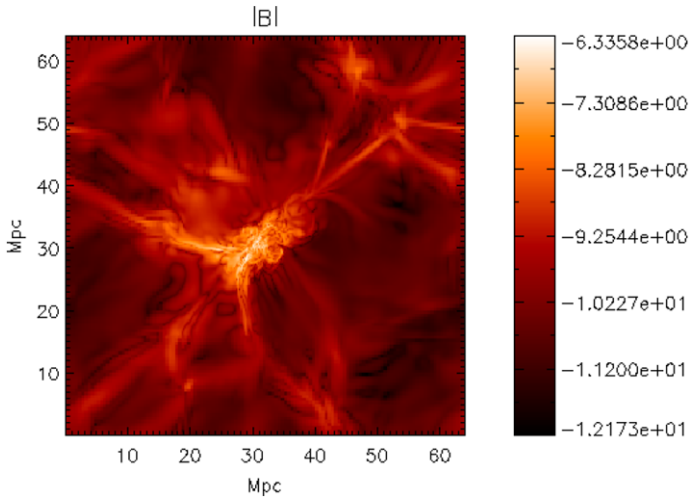


Fig. 9 Slice through the centre of the simulated box hosting a massive galaxy cluster. The simulation used the AMR code *FLASH* and followed the evolution of a weak magnetic field seed. The figure shows the logarithm of the magnetic field strength, in Gauss, measured at $z = 0.5$. From Brüggen et al. (2005)

semi-analytic simulations of magnetised galactic winds coupled to high-resolution N-body simulations of structure formation to estimate limits for the fraction of the ICM which can be significantly magnetised. Interestingly, the fraction of the volume of the ICM, which is predicted to be magnetised at a significant level by the galactic outflows during structure formation, is very high. In fact, the volume filling is high enough to be compatible with observations and a moderately efficient turbulent dynamo, operating within the ICM, will easily amplify these fields up to the observed values.

Additional to any dynamo action, further support for strong magnetic field amplification during the process of structure formation comes from the application of the Zel'dovich approximation (Zel'dovich 1970) to follow the MHD equations during the gravitational collapse (Bruni et al. 2003; King and Coles 2006). These papers demonstrate the existence of a super-adiabatic amplification due to the anisotropy of the collapse of the LSS within the cold dark matter paradigm.

A novel aspect of including magnetic field pressure into LSS simulations—even if in a simplified way—is investigated by Gazzola et al. (2007), which followed a purely empirical approach. Inspired by the scaling between magnetic field and gas density, found in both simulations and observations (Dolag et al. 2001), the magnetic field is parameterised as $B(z) = B_0(1+z)^{3\alpha}\rho^{\alpha}$. Then an isotropic pressure ($\propto B^2$) term is added into the equation of motion. This approach allows a quite effective way to identify which models of magnetic fields, along with their parameters, provide a non-thermal pressure support which is able to modify significantly the formation of structures. Thereby this approach gives a unique insight for the possible consequences of certain magnetic field scenarios.

5 Cosmic Rays

5.1 Modelling of Cosmic Rays (CRs)

CRs may originate in a number of cosmological processes including jets due to mass ac-

cretion, large-scale shocks and supernova remnants (see Sect. 2). One can conventionally distinguish two important classes of CR modelling in cosmological simulations.

The first one is related to the direct modelling of the origin, spatial distribution and spectral evolution of energetic electrons and positrons, which are responsible for most of the non-thermal emission observed so far in cosmological objects. In the vast majority of the non-thermal lepton ($e^{\pm}$) simulations, $e^{\pm}$ are considered as test particles scattered and accelerated by electro-magnetic fields associated with the cosmological flows. In the highly relativistic energy regime, the leptons radiate very efficiently and therefore suffer from radiative losses. Most of the models are based on the linear kinetic equations for the electron distribution function without accounting for the back-reaction that the accelerated electrons have on the bulk plasma motions and on the MHD-turbulence.

The second class of the CR simulations is related to the acceleration and propagation of energetic baryons. The radiation inferred from the energetic trans-relativistic baryons is basically caused by their nuclear interactions with the ambient matter which produces γ-rays, neutrino and $e^{\pm}$-pairs. Non-relativistic super-thermal baryons can produce some emission by inverse-Bremsstrahlung mechanisms. The most important feature of that class of processes, however, is that energetic non-thermal baryons could absorb a substantial fraction of the ram pressure of both non-relativistic and relativistic plasma flows. The back-reaction effects of energetic non-thermal particles on the plasma flows must then be taken into account. The most popular example of such a non-linear process is the diffusive particle acceleration by astrophysical shocks (see e.g. Blandford and Eichler 1987; Malkov and Drury 2001, for a review). The process of diffusive particle acceleration by MHD shocks could be very efficient in highly turbulent cosmic plasma.

In Fig. 1 of Sect. 2 we illustrated the simulated efficiency of the conversion into CRs of the power of shocks and of the bulk plasma motions in the particular cosmological context which represents the typical environment within galaxy clusters. The non-linear effect of the back-reaction of accelerated particles on large-scale plasma flows results in the specific temporal evolution of the particle spectra shown in Fig. 2 of Sect. 2. Moreover, as it was mentioned above, under a certain condition, thanks to the conversion of a fraction of the shock ram pressure into magnetic field energy, an efficient acceleration of baryons by MHD shocks in a turbulent cosmic plasma results in a strong amplification of the magnetic field in the shock upstream.

The acceleration of relativistic electrons and their emission were modelled in different cosmological contexts. To study the gradients of the spectral indices inferred from the radio observations at 1.4–5.0 GHz of five luminous 3C radio galaxies, Myers and Spangler (1985) developed a model in which the radiation is due to an isotropic ensemble of relativistic electrons which are subject to synchrotron radiation losses (synchrotron aging). Myers and Spangler (1985) inferred the age of the electrons, namely the time since their acceleration, at various locations in the lobes, and consequently derived the speed of the separation of the hot spots from the lobe material. The inferred speeds were found to be in the range 10 000–30 000 km s^{-1}. These results are consistent with the beam model, in which the lobe material is left behind by a hot spot advancing through the intergalactic medium at speeds of about 10 000 km s^{-1}.

Enßlin and Gopal-Krishna (2001) modelled different scenarios for the evolution of the radio plasma inside the cocoons of radio galaxies, after the activity of the central engine has ceased. These authors discussed an analytical model for the evolution of a relativistic electron population under synchrotron, inverse Compton and adiabatic energy losses or gains. It was demonstrated that fossil radio plasma with an age of up to 2 Gyr can be revived by compression in a shock wave of large-scale structure formation, caused during the merger

events or the mass accretion of galaxy clusters. The scenario was applied to explain the origin of the diffuse radio emission found in clusters of galaxies, without any likely parent radio galaxy seen nearby. The model predicts the existence of a population of diffuse radio sources, emitting at very low frequencies with an ultra-steep spectrum, which are located inside and possibly outside galaxy clusters. Thereby, extended radio emission could be interpreted as tracing the revival of aged fossil radio plasma which is caused by the shock waves associated with large-scale structure formation.

Some processes related to CRs were implemented in cosmological simulation codes. For instance, *COSMOCR* is a numerical code for the investigation of CRs in computational cosmology (Miniati 2001). The code includes a number of prescriptions to account for the diffusive shock acceleration, the mechanical and radiative energy losses and the spatial transport of the energetic particles into the cosmic environment. Primary CR electrons and ions are injected at the shock sites according to the phenomenological thermal leakage prescription. Secondary electrons are continuously injected as a results of proton-proton (p-p) inelastic collisions of primary CR ions and thermal background nuclei. The code consists of a conservative, finite volume method with a power-law sub-grid model in momentum space. Note that in this numerical approach, the back-reaction of the non-thermal components (CRs and magnetic fields) caused by their pressure contribution to the thermal gas is neglected.

To study the impact of CRs on galaxy and cosmic structure formation and evolution, Enßlin et al. (2007) developed an approximative framework which treats dynamical and radiative effects of CRs in cosmological simulations. These authors approximate the CR spectrum of each fluid element by a single power-law, with spatially and temporally varying normalisation, low-energy cut-off, and spectral index. In this framework they included some approximate prescriptions for CR injection and acceleration by shocks, as well as CR transport and energy losses due to Coulomb interactions, ionisation losses, Bremsstrahlung losses, and hadronic interactions with the background matter. These prescriptions are suited to be included into the global schemes of the numerical simulations of galaxy and structure formation. Although in such implementation the description of the CR population is more simplistic than in the work described earlier, the dynamical influence of the CRs onto the underlying hydrodynamics is no longer neglected. This is not only important for the dynamics of the ICM itself but also for the injection of the CRs by shocks, which are altered by the presence of the non-thermal pressure support of the CRs.

5.2 The Quest for Radio Relics

Roettiger et al. (1999a) were able to reproduce the main features of the extended peripheral radio emission (the so-called radio relics) observed in A 3667 by combining the single merger simulations (Sect. 4.1) with a model for the in situ re-acceleration of the relativistic particles, as described in the previous section. In their models, they injected, into the ICM, relativistic electrons with a power-law spectrum, where the power-law index $\gamma = 3/(r-1)+1$ is related to the gas compression ratio $r = \rho_2/\rho_1$ at the shock. They also related the age of the radio plasma t_a to the distance $d = \kappa v_s t_a$, using a weak-field/high-diffusion limit $\kappa = 1$. By having effective shock velocities $v_s \approx 700$–$1000\ \mathrm{km\,s^{-1}}$ and by aging the synchrotron spectrum with the formalism of Myers and Spangler (1985), they were able to reproduce the observed distribution of the spectral index for a magnetic field of $\approx 0.6\ \mu\mathrm{G}$ at the position of the radio relic (Fig. 10). Since such configurations seem to be quite common in galaxy clusters, it naturally raises the question, why not all the clusters show such peripheral radio emission. One possible explanation is that such shock structures are relatively short lived compared to the merger event itself. Also, it is necessary to have the

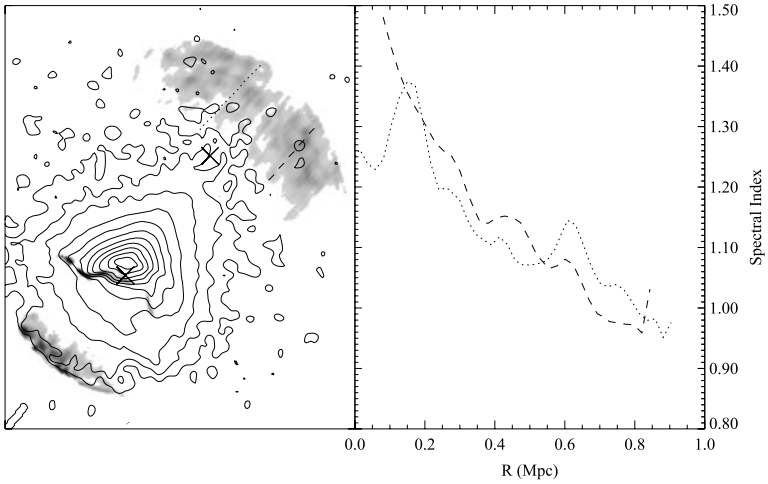


Fig. 10 Results from a simulation intended to reproduce the merging system observed in A 3667. The *left panel* shows the simulated X-ray surface brightness (*contours*) and the radio emission at 1.4 GHz (grey scale). The image is 3.15 × 3.85 Mpc. The *dashed* and *dotted lines* refer to the location of the radio spectral index ($\alpha_{4.9}^{1.4}$) profiles displayed in the *right panel*. From Roettiger et al. (1999a)

presence of a large-scale magnetic field. It could further be that only massive clusters can provide enough magnetic field and strong enough merger events to trigger such a peripheral emission.

To overcome this problem, Enßlin and Brüggen (2002) proposed that such radio relics could be made by a pre-existing fossil radio plasma illuminated by the shock waves that the merger events originate. In their models, following Enßlin and Gopal-Krishna (2001) which take into account synchrotron, inverse Compton and adiabatic energy losses and gains, they evolved the electron spectrum for the tracer particles, which represents the fossil radio plasma. Their simulation, using the *ZEUS* code, follows the evolution of a sphere of tracer particles hit by a shock front (Fig. 11). Such a configuration nicely reproduces the filamentary radio emission and toroidal structures observed in many cases. These simulations also predict that the magnetic fields are mostly aligned with the direction of the filaments, as suggested by observational data.

Hoeft et al. (2004) investigated this idea in a more realistic modelling by using the SPH code *GADGET* (Springel et al. 2001) to simulate a merging galaxy cluster within a cosmological environment. Such a simulation showed that the probability for a shock wave to flare the radio plasma is highly suppressed in the central regions of galaxy clusters, unlike the peripheral regions, where they found illuminated structures of size up to Mpc scales (Fig. 12). The reason for this is that the radio plasma ages much faster in the cluster centre than in the outer regions. In fact, in the cluster centre, the pressure of the radio plasma is higher and its energy losses due to the higher magnetic field are larger; moreover, the compression ratio of the shock wave is much higher in the low-density peripheral regions than in the cluster centre.

It is worth noting that a necessary condition to form such relics is that the initial state of the fossil radio plasma is characterised by a ratio of the magnetic to thermal pressure, $P_{\rm B}/P_{\rm gas}$, which has to be as low as 1 per cent to allow shocks to revive ≈ 1 Gyr old radio ghosts. It is also important to mention that Hoeft et al. (2004) find a high probability of radio emission, outside the shocks, due to the drained gas flows induced by the merger events,

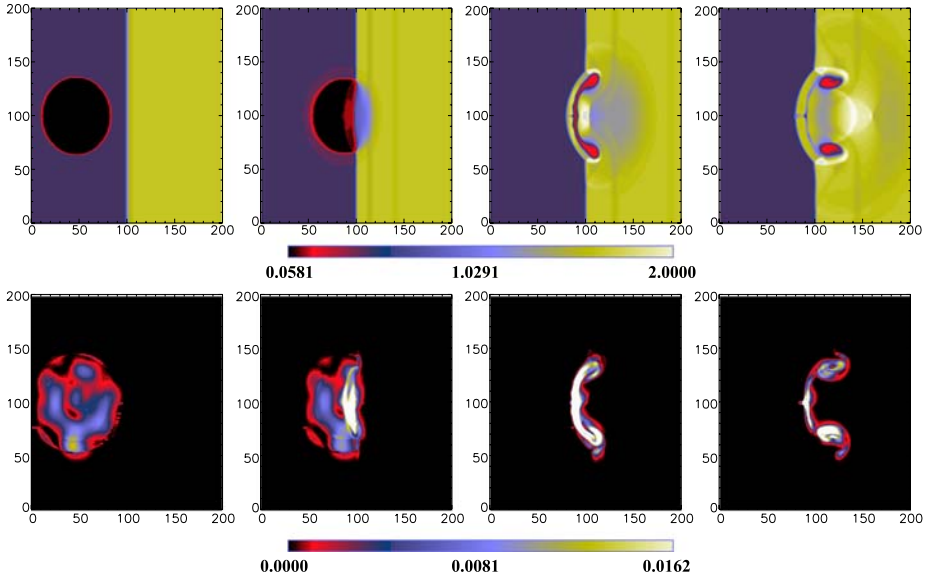


Fig. 11 Results from a simulation of a shock interacting with a fossil radio plasma. The *upper* (*lower*) *panels* show the time evolution of the gas density (magnetic field energy density) during the interaction. From Enßlin and Brüggen (2002)

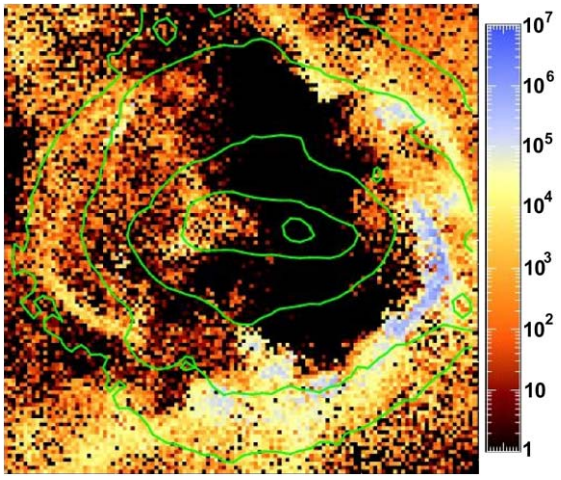


Fig. 12 Results from a simulation of a galaxy cluster undergoing a major merger. The figure shows the projected 'potential' radio luminosities for 1.13 Gyr old radio plasma, where $P_B/P_{\rm gas} = 0.01$. For comparison, the bolometric surface X-ray luminosity (*contours*) is given. The total bolometric X-ray luminosity of the cluster is $2 \times 10^{44}\,{\rm erg\,s^{-1}}$ and the emission-weighted temperature is 3 keV. Adapted from Hoeft et al. (2004)

which transport material from the outskirts towards the higher density regions. Therefore in some cases the adiabatic compression seems to be sufficient to revive the fossil radio plasma.

Recently, the picture of diffuse radio emission of galaxy clusters has changed, as more and more detailed observations indicate a more complex picture than the two main phenomena of radio halos and relics (see Ferrari et al. 2008—Chap. 6, this issue). It can well be that both scenarios described above are at work in galaxy clusters. The in situ re-acceleration of the relativistic particles at shocks might be responsible for the large, external relics, whereas

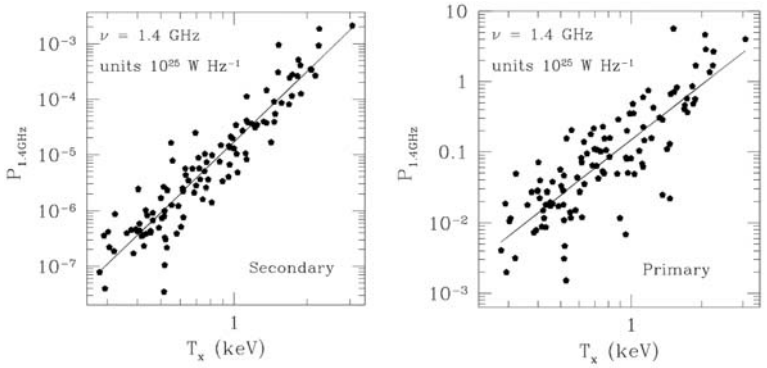


Fig. 13 Predicted radio properties of halos within a cosmological simulation of a cubic volume with a side length of 50 Mpc/h, resolved with 256^3 grid cells. The *left panel* shows the synchrotron power of these halos at 1.4 GHz from secondary electrons. The *right panel* shows the synchrotron power from primary electrons, mostly located in the peripheral regions. Note that the values for the luminosities due to the primary electrons should be scaled with the electron to proton injection ratio $R_{\mathrm{e/p}}$. From Miniati et al. (2001)

the compression of the fossil radio plasma might lead to the smaller relics which are more disturbed in shape and are usually found closer to the cluster centre.

5.3 The Quest for Radio Haloes

The accretion shocks from the LSS formation are also the location of the acceleration of the CRs that will be accreted into the LSS, specially within galaxy clusters. Using the *COSMOCR* code (Miniati 2001), Miniati et al. (2001) followed primary ions and electrons (injected and accelerated by diffuse cosmic shocks) and secondary electrons and positrons (produced in p-p inelastic collisions of CR ions with thermal ICM nuclei) within a cosmological simulation. Under the assumption that the magnetic field produced by the battery effect reflects a fair representation of the true distribution of the relative magnetic field strengths within the LSS, they were able to predict the central radio emission (radio halos), mainly produced by secondary electrons in a self-consistent treatment (Fig. 13).

Within their modelling they also reproduced the peripheral radio emission (radio relics), mainly produced by primary CRs, where the resulting morphology, polarisation and spectral index match the observed counterparts. However, one has to note that the extrapolation of these simulations (which are on group scale) to the observed data (which are on the scale of massive clusters) might not be straightforward (Fig. 14). It is further worth commenting that even with an electron to proton injection ratio $R_{\mathrm{e/p}}$ of 10^{-2} (which is derived from observations, see Table 5 in Miniati et al. 2001), the predicted luminosities for the primary emission are still significantly higher than for the secondary one (Fig. 14); this prediction does not seem to be confirmed by the observations. Although these results based on a secondary model generally agree with previous work based on simulations of massive galaxy clusters (Dolag and Enßlin 2000), they highlight the general problem that one faces when it is applied the secondary model for relativistic electrons. This model predicts radio haloes for every galaxy cluster, which is in strong contradiction with current observations, as summarised in Fig. 14, which also includes the upper limits obtained by a recent observation campaign at the GRMT (see Brunetti et al. 2007).

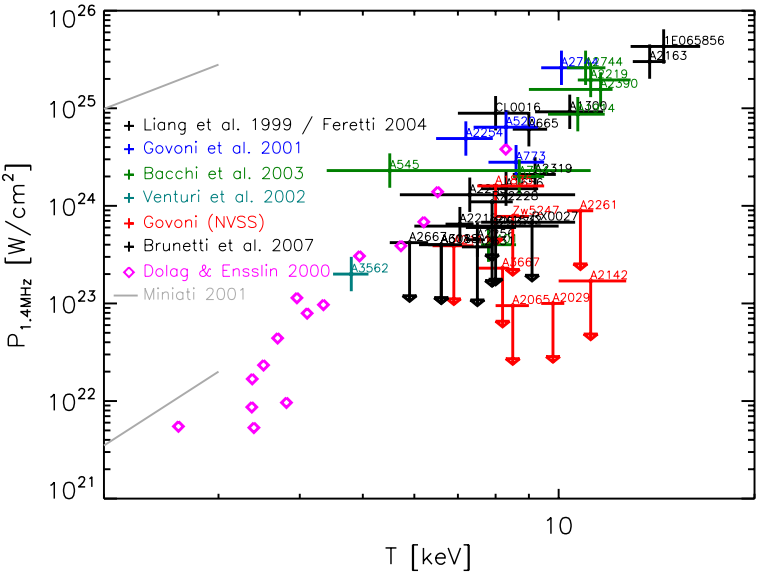


Fig. 14 Total power of radio halos observed at 1.4 MHz vs. cluster temperature. We plot the data from Liang et al. (2000), which were partially re-observed by Feretti (2007, in preparation) together with data from Govoni et al. (2001); Bacchi et al. (2003); Venturi et al. (2003). Some additional upper limits are collected thanks to the help of F. Govoni. Additional upper limits from Brunetti et al. (2007) are shown. We applied a secondary hadronic model, as described in Dolag and Enßlin (2000), to calculate the radio emission from the simulated galaxy clusters. We also added the predictions for the emission from primary (*upper line*) and secondary (*lower line*) electrons taken from Miniati et al. (2001) within the temperature range covered by the simulation. Note that the values for the luminosities for primary electrons should be scaled with the electron to proton injection ratio $R_{e/p}$

6 Turbulence

6.1 Simulations

During the growth of the LSS, galaxy clusters continuously accrete other structures, most of the time smaller structures (like galaxies), but sometime also objects with similar mass (major mergers). All these structures continue to move with supersonic or transsonic velocities inside the cluster. Together with the diffuse accretion (which is generally anisotropic) and the generation of turbulence by hydrodynamic instabilities induced by these bulk motions, the gas in clusters of galaxies generally contains an amount of kinetic energy which is not negligible compared to the amount of thermal energy.

It is worth mentioning that different simulation methods reach good agreement in predicting that the ratio of bulk kinetic energy to thermal energy is up to 15% in galaxy clusters (see Frenk et al. 1999). Based on cosmological hydrodynamic simulations, Inogamov and Sunyaev (2003) pointed out that the broadening of the emission lines in the X-ray band (e.g. the iron K line) due to these expected bulk motions is appreciably larger than the broadening due to thermal motions. Thus, instruments like the ones designed for Suzaku or future instruments like XEUS will be able to infer such bulk motions from the analysis of the line shapes.

Even after subtraction of the cluster peculiar velocity, large-scale bulk motions will be the main contributor to the deformation of the line shapes. The imprint of turbulence induced

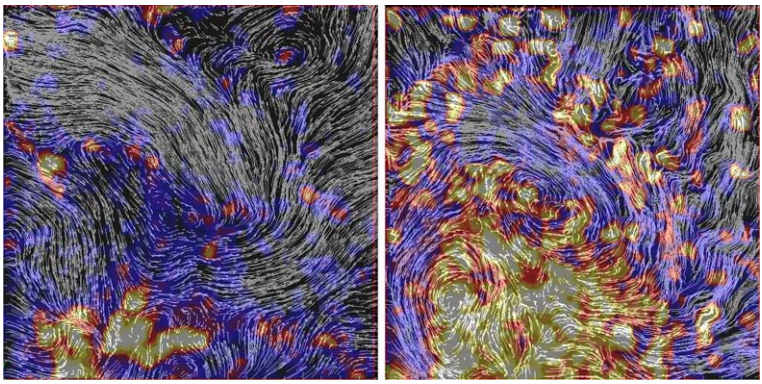


Fig. 15 The gas velocity field in a slice through the central Mpc region of a simulated cluster after subtracting the *global* mean bulk velocity of the cluster. The *left panel* shows a run with the original SPH viscosity, the *right panel* for a low-viscosity scheme. The underlying colour maps represent the ratio of turbulent kinetic energy to total kinetic energy content of the particles, inferred after subtracting the *local* mean velocity, as described in Dolag et al. (2005b)

by hydrodynamic instabilities along flows (e.g. shear flows) might instead be more subtle to infer from the line profiles.

In recent high-resolution SPH simulations of galaxy clusters within a cosmological environment, Dolag et al. (2005b) used a novel scheme to treat artificial viscosity within the *GADGET2* code (Springel 2005) to quantify how intense the shear flow must be to drive fluid instabilities. Being a quite common process within cosmic structure formation, these fluid instabilities driven by shear flows strongly increase the level of turbulence within the ICM (Fig. 15).

Although this small-scale turbulence contributes a measurable effect to the non-thermal pressure within the ICM, the effect on the line shapes is much smaller than the broadening caused by the large-scale bulk motions. Applying this method to a set of simulated galaxy clusters, Vazza et al. (2006) showed that for relaxed clusters the turbulent energy content in the ICM scales with the thermal one. The fraction of turbulent to thermal energy turned out to be rather independent of the mass of the system. Interestingly, this is consistent with the semi-analytical predictions of Cassano and Brunetti (2005), who estimated the amount of turbulent energy within the ICM in terms of the $P\mathrm{d}V$ work done by the infalling substructures.

6.2 Observational Evidence

The spectroscopic resolution of current X-ray instruments is not high enough to measure the line shapes and infer the level of turbulence present within the ICM. However, using a mosaic of XMM-Newton observations of the Coma cluster, Schuecker et al. (2004) were able to produce spatially-resolved gas pressure maps which indicate the presence of a significant amount of turbulence. Performing a Fourier analysis of the data reveals the presence of a scale-invariant pressure fluctuation spectrum in the range between 40 and 90 kpc which is well described by a projected Kolmogorov turbulence spectrum (see Fig. 16); at least 10 percent of the total ICM pressure is in turbulent form.

Alternatively, in a more indirect way, Enßlin and Vogt (2006) argue that the recently reported Kolmogorov-like magnetic turbulence spectrum, which is inferred from Faraday

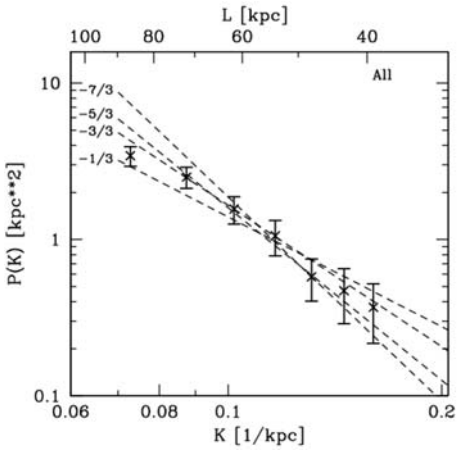


Fig. 16 The dots with error bars are the observed projected power spectral densities as inferred from X-ray observations of the Coma cluster, after subtraction of the shot noise; the *dashed lines* are model predictions. From Schuecker et al. (2004)

rotation measurements in the cool core of clusters, can be understood by kinetic energy injection (see also Subramanian et al. 2006). Such a dynamo model predicts the correct magnetic field strength in cool core clusters for reasonable (but yet not directly observable) values of the hydrodynamic turbulence velocity and characteristic length scales. This result indicates that the magnetic fields might directly reflect the presence of hydrodynamic turbulence. Such models are directly related to the turbulence induced by buoyant radio lobes from the central radio galaxy which rise into the cluster atmosphere. On the other hand, Faraday rotation within non-cooling flow clusters with multiple extended radio sources (which therefore probe the magnetic field structure at different radii) can give alternative constraints on the magnetic field power spectra (see Murgia et al. 2004; Govoni et al. 2006) and thus on the underlying hydrodynamic turbulence present in the ICM.

7 Concluding Remarks

It seems that, in the last years, a consistent picture of magnetic fields in clusters of galaxies has emerged from both numerical work and observations. Simulations of individual processes like shear flows, shock/bubble interactions or turbulence/merging events predict a super-adiabatic amplification of magnetic fields. It is worth mentioning that this common result is obtained by using a variety of different codes, which are based on different numerical schemes. Further support for such super-adiabatic amplifications comes from analytical estimates of the anisotropic collapse which use the Zel'dovich approximation (Zel'dovich 1970). When this amplification occurs in fully consistent cosmological simulations, various observational aspects are reproduced (e.g. Fig. 8); moreover, the final strength of the magnetic field reaches a level sufficient to link models that predict magnetic field seed at high redshift with the magnetic fields observed in galaxy clusters today. It is important to note that all the simulations show that this effect increases when the resolution is improved, and therefore all the current values have to be taken as lower limits of the possible amplification.

Despite this general agreement, there are significant differences among the predictions on the structure of the magnetic fields coming from different models of magnetic field seeds. In particular, there are several differences between the up-scaled cosmological battery fields (Miniati et al. 2001; Sigl et al. 2004) and the magnetic field predicted by the high-resolution

simulations of galaxy clusters that use either AMR (Brüggen et al. 2005) or SPH (Dolag et al. 1999, 2002, 2005a). In the SPH case, it is possible to follow the amplification of the field seeds within the turbulent ICM in more detail.

A good visual impression can be obtained by comparing the regions filled with the high magnetic fields shown in Figs. 6 and 9. It is clear that the high magnetic field regions for the battery fields are predicted to be much more extended; this result leads to a flat profile around the forming structure, whereas for the turbulent amplified magnetic fields the clusters show a magnetic field distribution which is much more peaked. Part of this difference originates from the physical model, as the cosmological shocks are much stronger outside the clusters. Somewhat less clear is how big it is the influence of the different numerical resolutions of these simulations to these discrepancies. Usually, the model parameters for such simulations can only be calibrated using the magnetic field strength and structure within the high-density regions of galaxy clusters. Therefore it is crucial to perform detailed comparisons with all the available observations to validate the simulations. Note that extrapolating the predictions of the simulations into lower density regions, where no strong observational constraints exist, will further amplify the differences among the simulations in the predictions of the magnetic field structure.

Moreover, one has to keep in mind that, depending on the ICM resistivity, the magnetic field could suffer a decay, which is so far neglected in all the simulations. Furthermore, a clear deficiency of the current simulations is that they do not include the creation of a magnetic field due to all the feedback processes happening within the LSS (like radio bubbles inflated by AGN, galactic winds, etc.): this might alter the magnetic field prediction if their contribution turns out to be significant. Also all of the simulations so far neglect radiative losses; if included, this would lead to a significant increase of the density in the central region of clusters and thereby to a further magnetic field amplification in these regions. Finally, there is an increasing number of arguments suggesting that instabilities and turbulence on very small scales can amplify the magnetic fields to the observed μG level over a relatively short timescale. However, it is still unclear how such fields, which are tangled on very small scales, can be further processed to be aligned on large scales (up to hundreds of kpc), as observed.

Concerning radio haloes and relics, the picture is only partially consistent (for a more detailed discussion of primary and secondary models see Brunetti (2004) and references therein as well as Ferrari et al. 2008—Chap. 6, this issue). Radio relics seem to be most likely related to strong shocks produced by major merger events and therefore produced by direct re-acceleration of CRs, the so-called primary models. Although some of the observed features, like morphology, polarisation and position with respect to the cluster centre, can be reasonably well reproduced, there might be still some puzzles to be solved. On the one hand, direct acceleration of CRs in shocks seems to overestimate the abundance, and maybe the luminosity, of radio relics; on the other hand, simulations which illuminate fossil radio plasma can produce reasonable relics only starting from a small range of parameter settings.

A similar situation arises for modelling the central radio emission of galaxy clusters by secondary models. The total luminosity seems to be reproduced using reasonable assumptions and also the observed steep correlation between cluster temperature/mass and radio power seems to be reproduced quite well. However, these models suffer from two drawbacks. The first one is that, in the framework of such models, every massive cluster produces a powerful radio halo, at odds with observations (Fig. 14). The other problem of secondary models is that the detailed radio properties are not reproduced. In fact, firstly, in most cases, the profile of radio emission is too steep, so that these models can almost never reproduce the size of the observed radio halos. Secondly, the observed spectral steepening (e.g. Giacintucci et al. 2005) cannot be reproduced.

One possible way out, at least for the first concern, would be that magnetic fields are extremely dynamical and transient and therefore *light up* the halo during a merger. However, there is also no indication from observations that clusters which show radio emission contain magnetic fields which are more intense than those in clusters without observable extended diffuse radio emission. On the contrary, the cluster A 2142 has a magnetic field strength similar to the Coma cluster, but the upper limit on its radio emission is at least two orders of magnitude below the value expected from the correlation (Fig. 14). Note that both clusters are merging systems which are characterised by the presence of two central cD galaxies. This indicates that there should be further processes involved or additional conditions required to produce radio emission. It is worth noting that recent models, based on turbulent acceleration, seem to overcome this problem (see Kuo et al. 2003; Brunetti and Blasi 2005; Cassano and Brunetti 2005, and references therein). Cassano and Brunetti (2005) predicted that the probability for a galaxy cluster to show a giant radio halo is an increasing function of the cluster mass; these authors also reproduced the observed probability of ≈ 30 % for a massive galaxy clusters to show extended radio emission.

Being dominated by large-scale bulk motions, induced by the cosmological structure formation process, different simulation methods reach good agreement in predicting that the ratio of the bulk kinetic energy to thermal energy has an upper limit of 15% in galaxy clusters (see Frenk et al. 1999). However, the exact role of turbulence induced by hydrodynamic instabilities within the complex flow patterns of galaxy clusters is still hard to quantify (see Dolag et al. 2005b). It should also be mentioned that the real viscosity of the ICM is hardly known, and can play a noticeable role within the ICM, as demonstrated in simulations which attempt to include physically motivated viscosity into a cosmological context (see Sijacki and Springel 2006).

The dynamical role of CRs in galaxy clusters is not very well understood yet. Simulations using different codes predict quite different relative pressure contained in CRs within galaxy clusters. Moreover, the simulations indicate that the relative importance of CRs in galaxy clusters also strongly depends on other non-thermal processes, like radiative losses and feedback from star formation (see Pfrommer et al. 2007).

In general it has to be pointed out that, although the complexity of the physical processes treated by cosmological simulations have improved dramatically in recent years, this research field is still young. Therefore future work is expected to enlighten such complex processes within the ICM and will help to improve our understanding of the non-thermal components in galaxy clusters. Such work will also be needed to interpret the observational information which is overwhelmingly increasing in quality and size and is expected to become available soon with the next generation of instruments, particularly at radio band, like ALMA, LOFAR, SKA, EVLA and others.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". Special thanks to the anonymous referee for suggestions that improved the manuscript, Torsten Enßlin for various helpful discussions and Daniel Clarke for carefully reading the manuscript. A.M.B. acknowledges the support from RBRF grant 06-02-16844 and a support from RAS Programs. A.D. also gratefully acknowledges partial support from the PRIN2006 grant "Costituenti fondamentali dell'Universo" of the Italian Ministry of University and Scientific Research and from the INFN grant PD51.

References

M. Bacchi, L. Feretti, G. Giovannini, F. Govoni, Astron. Astrophys. **400**, 465 (2003)
A.R. Bell, S.G. Lucek, Mon. Not. R. Astron. Soc. **321**, 433 (2001)
V.S. Berezinsky, P. Blasi, V.S. Ptuskin, Astrophys. J. **487**, 529 (1997)

S. Bertone, C. Vogt, T. Enßlin, Mon. Not. R. Astron. Soc. **370**, 319 (2006)
J. Binney, G. Tabor, Mon. Not. R. Astron. Soc. **276**, 663 (1995)
G.T. Birk, H. Wiechen, A. Otto, Astrophys. J. **518**, 177 (1999)
D. Biskamp, *Magnetohydrodynamic Turbulence* (Cambridge University Press, Cambridge, 2003)
R. Blandford, D. Eichler, Phys. Rep. **154**, 1 (1987)
M. Brüggen, M. Ruszkowski, A. Simionescu, M. Hoeft, C. Dalla Vecchia, Astrophys. J. **631**, L21 (2005)
G. Brunetti, J. Korean Astron. Soc. **37**, 493 (2004)
G. Brunetti, P. Blasi, Mon. Not. R. Astron. Soc. **363**, 1173 (2005)
G. Brunetti, T. Venturi, D. Dallacasa et al., Astrophys. J. **670**, L5 (2007)
M. Bruni, R. Maartens, C.G. Tsagas, Mon. Not. R. Astron. Soc. **338**, 785 (2003)
A.M. Bykov, Space Sci. Rev. **99**, 317 (2001)
A. Bykov, K. Dolag, F. Durret, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9312-9
R. Cassano, G. Brunetti, Mon. Not. R. Astron. Soc. **357**, 1313 (2005)
E. Churazov, W. Forman, C. Jones, H. Böhringer, Astrophys. J. **590**, 225 (2003)
T.E. Clarke, P.P. Kronberg, H. Böhringer, Astrophys. J. **547**, L111 (2001)
K. Dolag, M. Bartelmann, H. Lesch, Astron. Astrophys. **348**, 351 (1999)
K. Dolag, M. Bartelmann, H. Lesch, Astron. Astrophys. **387**, 383 (2002)
K. Dolag, T.A. Enßlin, Astron. Astrophys. **362**, 151 (2000)
K. Dolag, S. Schindler, F. Govoni, L. Feretti, Astron. Astrophys. **378**, 777 (2001)
K. Dolag, D. Grasso, V. Springel, I. Tkachev, J. Exp. Theor. Phys. Lett. **79**, 583 (2004)
K. Dolag, D. Grasso, V. Springel, I. Tkachev, J. Cosm. Astro-Part. Phys. **1**, 9 (2005a)
K. Dolag, F. Vazza, G. Brunetti, G. Tormen, Mon. Not. R. Astron. Soc. **364**, 753 (2005b)
T.A. Enßlin, P.L. Biermann, P.P. Kronberg, X.-P. Wu, Astrophys. J. **477**, 560 (1997)
T.A. Enßlin, M. Brüggen, Mon. Not. R. Astron. Soc. **331**, 1011 (2002)
T.A. Enßlin, Gopal-Krishna, in *Particles and Fields in Radio Galaxies Conference*. ASP Conf. Ser., vol. 250 (2001), pp. 454
T.A. Enßlin, C. Vogt, Astron. Astrophys. **453**, 447 (2006)
T.A. Enßlin, C. Pfrommer, V. Springel, M. Jubelgas, Astron. Astrophys. **473**, 41 (2007)
L. Feretti, D. Dallacasa, G. Giovannini, A. Tagliani, Astron. Astrophys. **302**, 680 (1995)
L. Feretti, D. Dallacasa, F. Govoni et al., Astron. Astrophys. **344**, 472 (1999)
C. Ferrari, F. Govoni, S. Schindler, A. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9311-x
C.S. Frenk, S.D.M. White, P. Bode et al., Astrophys. J. **525**, 554 (1999)
S.R. Furlanetto, A. Loeb, Astrophys. J. **556**, 619 (2001)
S. Gabici, P. Blasi, Astrophys. J. **583**, 695 (2003)
L. Gazzola, E.J. King, F.R. Pearce, P. Coles, Mon. Not. R. Astron. Soc. **375**, 657 (2007)
S. Giacintucci, T. Venturi, G. Brunetti et al., Astron. Astrophys. **440**, 867 (2005)
M. Gitti, B.R. McNamara, P.E.J. Nulsen, M.W. Wise, Astrophys. J. **660**, 1118 (2007)
F. Govoni, T.A. Enßlin, L. Feretti, G. Giovannini, Astron. Astrophys. **369**, 441 (2001)
F. Govoni, M. Murgia, L. Feretti et al., Astron. Astrophys. **460**, 425 (2006)
D. Grasso, H.R. Rubinstein, Phys. Rep. **348**, 163 (2001)
G. Gregori, F. Miniati, D. Ryu, T.W. Jones, Astrophys. J. **543**, 775 (2000)
M. Hoeft, M. Brüggen, G. Yepes, Mon. Not. R. Astron. Soc. **347**, 389 (2004)
N.A. Inogamov, R.A. Sunyaev, Astron. Lett. **29**, 791 (2003)
T.W. Jones, D.S. De Young, Astrophys. J. **624**, 586 (2005)
H. Kang, D. Ryu, R. Cen, D. Song, Astrophys. J. **620**, 21 (2005)
H. Kang, D. Ryu, R. Cen, J.P. Ostriker, Astrophys. J. **669**, 729 (2007)
K.-T. Kim, P.P. Kronberg, P.C. Tribble, Astrophys. J. **379**, 80 (1991)
E.J. King, P. Coles, Mon. Not. R. Astron. Soc. **365**, 1288 (2006)
P.P. Kronberg, H. Lesch, U. Hopp, Astrophys. J. **511**, 56 (1999)
R.M. Kulsrud, R. Cen, J.P. Ostriker, D. Ryu, Astrophys. J. **480**, 481 (1997)
P.-H. Kuo, C.-Y. Hwang, W.-H. Ip, Astrophys. J. **594**, 732 (2003)
H. Liang, R.W. Hunstead, M. Birkinshaw, P. Andreani, Astrophys. J. **544**, 686 (2000)
M.A. Malkov, L.O.C. Drury, Rep. Prog. Phys. **64**, 429 (2001)
F. Miniati, Comput. Phys. Commun. **141**, 17 (2001)
F. Miniati, D. Ryu, H. Kang et al., Astrophys. J. **542**, 608 (2000)
F. Miniati, T.W. Jones, H. Kang, D. Ryu, Astrophys. J. **562**, 233 (2001)
M. Murgia, F. Govoni, L. Feretti et al., Astron. Astrophys. **424**, 429 (2004)
S.T. Myers, S.R. Spangler, Astrophys. J. **291**, 52 (1985)
C.A. Norman, D.B. Melrose, A. Achterberg, Astrophys. J. **454**, 60 (1995)
V. Pavlidou, B.D. Fields, Astrophys. J. **642**, 734 (2006)

C. Pfrommer, V. Springel, T.A. Enßlin, M. Jubelgas, Mon. Not. R. Astron. Soc. **367**, 113 (2006)
C. Pfrommer, T.A. Enßlin, V. Springel, M. Jubelgas, K. Dolag, Mon. Not. R. Astron. Soc. **378**, 385 (2007)
V. Quilis, J.M.A. Ibanez, D. Saez, Astrophys. J. **502**, 518 (1998)
M.J. Rees, M.C. Begelman, R.D. Blandford, E.S. Phinney, Nature **295**, 17 (1982)
Y. Rephaeli, J. Nevalainen, T. Ohashi, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9314-7
C.S. Reynolds, B. McKernan, A.C. Fabian, J.M. Stone, J.C. Vernaleo, Mon. Not. R. Astron. Soc. **357**, 242 (2005)
K. Roettiger, J.O. Burns, J.M. Stone, Astrophys. J. **518**, 603 (1999a)
K. Roettiger, J.M. Stone, J.O. Burns, Astrophys. J. **518**, 594 (1999b)
C. Rordorf, D. Grasso, K. Dolag, Astropart. Phys. **22**, 167 (2004)
D. Ryu, H. Kang, P.L. Biermann, Astron. Astrophys. **335**, 19 (1998)
D. Ryu, H. Kang, E. Hallman, T.W. Jones, Astrophys. J. **593**, 599 (2003)
S. Schindler, A. Diaferio, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9321-8
P. Schuecker, A. Finoguenov, F. Miniati, H. Böhringer, U.G. Briel, Astron. Astrophys. **426**, 387 (2004)
G. Sigl, F. Miniati, T.A. Enssßlin, Phys. Rev. D **70**, 043007 (2004)
D. Sijacki, V. Springel, Mon. Not. R. Astron. Soc. **371**, 1025 (2006)
V. Springel, Mon. Not. R. Astron. Soc. **364**, 1105 (2005)
V. Springel, N. Yoshida, S. White, New Astron. **6**, 79 (2001)
J.M. Stone, M.L. Norman, Astrophys. J. Suppl. Ser. **80**, 753 (1992a)
J.M. Stone, M.L. Norman, Astrophys. J. Suppl. Ser. **80**, 791 (1992b)
K. Subramanian, A. Shukurov, N.E.L. Haugen, Mon. Not. R. Astron. Soc. **366**, 1437 (2006)
F. Vazza, G. Tormen, R. Cassano, G. Brunetti, K. Dolag, Mon. Not. R. Astron. Soc. **369**, L14 (2006)
T. Venturi, S. Bardelli, D. Dallacasa et al., Astron. Astrophys. **402**, 913 (2003)
A. Vladimirov, D.C. Ellison, A. Bykov, Astrophys. J. **652**, 1246 (2006)
H.J. Völk, A.M. Atoyan, Astrophys. J. **541**, 88 (2000)
H.J. Völk, F.A. Aharonian, D. Breitschwerdt, Space Sci. Rev. **75**, 279 (1996)
N. Werner, F. Durret, T. Ohasi, S. Schindler, R.P.C. Wiersma, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9320-9
Y.B. Zel'dovich, Astron. Astrophys. **5**, 84 (1970)

Chapter 16
Observations of Metals in the Intra-Cluster Medium

N. Werner · F. Durret · T. Ohashi · S. Schindler · R.P.C. Wiersma

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9320-9

Abstract Because of their deep gravitational potential wells, clusters of galaxies retain all the metals produced by the stellar populations of the member galaxies. Most of these metals reside in the hot plasma which dominates the baryon content of clusters. This makes them excellent laboratories for the study of the nucleosynthesis and chemical enrichment history of the Universe. Here we review the history, current possibilities and limitations of the abundance studies, and the present observational status of X-ray measurements of the chemical composition of the intra-cluster medium. We summarise the latest progress in using the abundance patterns in clusters to put constraints on theoretical models of supernovae and we show how cluster abundances provide new insights into the star-formation history of the Universe.

Keywords Galaxies: clusters · Galaxies: abundances · X-rays: galaxies: clusters

N. Werner (✉)
SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands
e-mail: n.werner@sron.nl

N. Werner
Max-Planck-Institute für Astrophysik, Karl-Schwarzschild-Strasse 1, 85749 Garching, Germany

F. Durret
Institut d'Atrophysique de Paris, CNRS, UMR 7095, Université Pierre et Marie Curie, 98bis Bd Arago, 75014 Paris, France

T. Ohashi
Institute of Space and Astronautical Science, JAXA, 3-1-1 Yoshinodai, Sagamihara, Kanagawa 229-8510, Japan

S. Schindler
Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria

R.P.C. Wiersma
Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands

1 Introduction

Clusters of galaxies are excellent astrophysical laboratories, which allow us to study the chemical enrichment history of the Universe. They have the deepest known gravitational potential wells which keep the metals produced in the stellar populations of the member galaxies within the clusters. About 70–90% of the baryonic mass content of clusters of galaxies is in the form of hot (10^7–10^8 K) X-ray emitting gas (Ettori and Fabian 1999). In this hot intra-cluster medium (ICM) the dominant fraction of cluster metals resides. To the extent that the stellar populations where the cluster metals were synthesised can be considered representative, the metal abundances in the ICM provide constraints on nucleosynthesis and on the star formation history of the Universe.

It is remarkable that all the abundant elements, that were synthesised in stars after the primordial nucleosynthesis, have the energies of their K- and L-shell transitions in the spectral band accessible to modern X-ray telescopes. Most of the observed emission lines in the ICM arise from the well understood hydrogen and helium like ions and their equivalent widths can be, under the reasonable assumption of collisional equilibrium, directly converted into the elemental abundance of the corresponding element. Complications arising from optical depth effects, depletion into dust grains, extinction, non-equilibrium ionisation are minimal or absent. Therefore, abundance determinations of the hot ICM are more robust than those of stellar systems, H II regions or planetary nebulae. This relatively uncomplicated physical environment makes the ICM an attractive tool for studies of the chemical enrichment (for the theoretical progress in metal enrichment processes see Schindler and Diaferio 2008—Chap. 17, this issue).

2 Sources of Metals

Most of the metals from O up to the Fe-group are produced by supernovae. The supernovae can be roughly divided into two groups: Type Ia supernovae (SN Ia) and core collapse supernovae (SN_{CC}).

SN Ia are most likely thermonuclear explosions of accreting white dwarfs. When the white dwarf reaches the Chandrasekhar limit, carbon ignition in the central region leads to a thermonuclear runaway. A flame front then propagates through the star at a subsonic speed as a deflagration wave. In the delayed-detonation models, the deflagration wave is assumed to be transformed into a detonation at a specific density. SN Ia produce a large amount of Fe, Ni, and Si-group elements (Si, S, Ar, and Ca). Contrary to SN_{CC}, they produce only very small amounts of O, Ne, and Mg. In Fig. 1 and Table 1 we show the theoretically calculated yields for different SN Ia models (Iwamoto et al. 1999). We show the yields of elements relative to the yield of Fe in the Solar units of Grevesse and Sauval (1998). Convective deflagration models are represented by the W7 and W70 models. The WDD1, WDD2, WDD3, CDD1, and CDD2 models refer to delayed-detonation explosion scenarios and the last digit indicates the deflagration to detonation density in units of 10^7 g cm^{-3}. The "C" and "W" refer to two different central densities (1.37×10^9 and 2.12×10^9 g cm^{-3}, respectively) in the model at the onset of the thermonuclear runaway. The relative yield of the Si-group elements can be a good indicator of the incompleteness of the Si-group burning into Fe-group elements. Delayed detonation models have typically higher yields of Si-group elements relative to Fe than the deflagration models. The Ni/Fe abundance ratio in the SN Ia ejecta is an indicator of neutron-rich isotope production, which depends on the electron capture efficiency in the core of the exploding white dwarf. The deflagration models produce larger Ni/Fe ratios than the delayed detonation models.

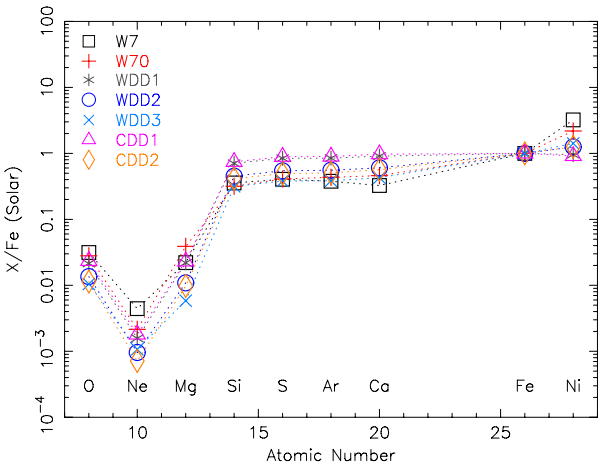


Fig. 1 Yields of elements relative to Fe in the Solar units of Grevesse and Sauval (1998) for different SN Ia models from Iwamoto et al. (1999)

Table 1 Yields of elements relative to Fe in the Solar units of Grevesse and Sauval (1998) for different SN Ia models from Iwamoto et al. (1999)

Element	W7	W70	WDD1	WDD2	WDD3	CDD1	CDD2
O	0.0314	0.0280	0.0214	0.0136	0.0105	0.0235	0.0114
Ne	0.00443	0.00215	0.00157	0.000960	0.00107	0.00180	0.000714
Mg	0.0221	0.0390	0.0219	0.0109	0.00584	0.0231	0.00976
Si	0.356	0.317	0.706	0.453	0.316	0.746	0.413
S	0.405	0.410	0.843	0.544	0.380	0.890	0.496
Ar	0.378	0.433	0.846	0.555	0.384	0.896	0.506
Ca	0.326	0.460	0.908	0.604	0.427	0.972	0.566
Fe	≡1.0	≡1.0	≡1.0	≡1.0	≡1.0	≡1.0	≡1.0
Ni	3.22	2.18	0.957	1.25	1.41	0.914	1.26

In the left panel of Fig. 2 and Table 2 we show the theoretical yields for three SN_{CC} models integrated over Salpeter and top-heavy initial mass functions (IMFs) between 10 and 50 $M_{\odot}$ (Woosley and Weaver 1995; Chieffi and Limongi 2004; Nomoto et al. 2006). See also Borgani et al. (2008—Chap. 18, this issue) for more details on stellar population models. In the right panel of Fig. 2 we show the SN_{CC} yields for the Salpeter IMF for different progenitor metallicities (where $Z = 0.02$ refers to Solar metallicity; Nomoto et al. 2006). The Fe yield of the model by Woosley and Weaver (1995) is about an order of magnitude lower than that in the models by Chieffi and Limongi (2004) and Nomoto et al. (2006). The plots also show that abundance ratios of O/Mg and Ne/Mg can be used to discriminate between different IMFs. The SN_{CC} yields for progenitors of different metallicities differ mainly in O/Ne/Mg ratio. Because the progenitor metallicities and perhaps also the IMF evolve, the applicability of comparing the observed data with a single average set of yields may be limited. However, the abundances of these elements in the ICM may still be used to infer to what extent was the *dominant fraction* of the SN_{CC} progenitors preenriched, and what was its IMF.

We note, that while the abundances in Figs. 1 and 2 are shown with respect to the Solar values of Grevesse and Sauval (1998), in literature and also in several figures of this

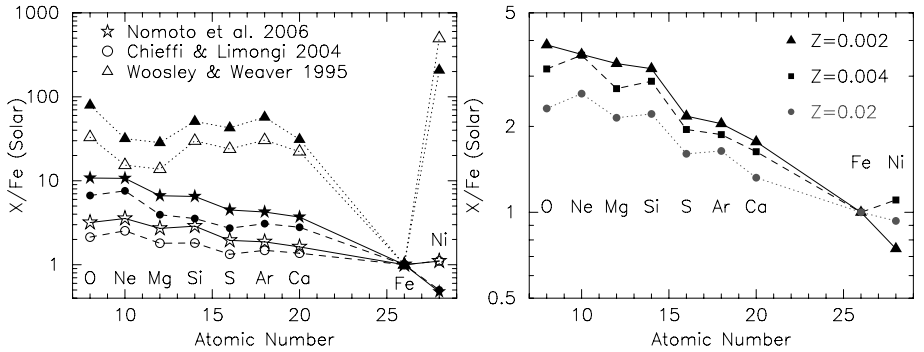


Fig. 2 *Left panel:* Yields of elements relative to Fe in the Solar units of Grevesse and Sauval (1998) for different SN_{CC} models with a progenitor metallicity of $Z = 0.004$ integrated over Salpeter and top-heavy initial mass functions between 10 and 50 $M_{\odot}$. *Right panel:* SN_{CC} yields for different progenitor metallicities ($Z = 0.02$ refers to Solar metallicity; Nomoto et al. 2006) integrated over the Salpeter IMF

Table 2 Yields of elements relative to Fe in the Solar units of Grevesse and Sauval (1998) for different SN_{CC} models with a progenitor metallicity of $Z = 0.004$ integrated over Salpeter and top-heavy initial mass functions between 10 and 50 $M_{\odot}$

Element	Nomoto et al. (2006)		Chieffi and Limongi (2004)		Woosley and Weaver (1995)	
	Salpeter	Top-heavy	Salpeter	Top-heavy	Salpeter	Top-heavy
O	3.18	10.8	2.13	6.67	33.0	79.6
Ne	3.55	10.7	2.53	7.59	15.4	31.8
Mg	2.71	6.63	1.80	3.94	13.8	28.5
Si	2.88	6.50	1.82	3.55	29.9	50.9
S	1.95	4.49	1.33	2.72	23.9	42.7
Ar	1.87	4.24	1.49	3.08	30.5	57.1
Ca	1.63	3.70	1.37	2.80	22.2	31.1
Fe	≡1.0	≡1.0	≡1.0	≡1.0	≡1.0	≡1.0
Ni	1.10	0.475	1.11	0.501	497	208

review the abundances are often shown with respect to the outdated Solar abundances by Anders and Grevesse (1989) or with respect to newer sets of Solar and proto-solar abundances by Lodders (2003). The more recent Solar abundance determinations of O, Ne, and Fe by Lodders (2003) are ∼30% lower than those given by Anders and Grevesse (1989). The Solar abundances of O and Ne reported by Grevesse and Sauval (1998) are higher and the abundance of Fe is slightly lower than those reported by Lodders (2003). A comparison of different Solar abundance values is shown in Table 3.

While elements from O up to the Fe-group are produced by supernovae, the main sources of carbon and nitrogen are still being debated. Both elements are believed to originate from a wide range of sources including winds of short-lived massive metal rich stars, longer-lived low- and intermediate-mass stars, and also an early generation of massive stars (e.g., Gustafsson et al. 1999; Chiappini et al. 2003; Meynet and Maeder 2002).

In the Galaxy, Shi et al. (2002) found that C is enriched by winds of metal-rich massive stars at the beginning of the Galactic disk evolution, while at a later stage it is produced mainly by low-mass stars. Bensby and Feltzing (2006) found that the C enrichment in the

Table 3 Comparison of Solar abundances by Anders and Grevesse (AG; 1989), Grevesse and Sauval (GS; 1998), and of the proto-solar abundances by Lodders (2003) on a logarithmic scale with H ≡ 12. These abundances are used at several places of this issue

Element	AG	GS	Lodders
H	12.00	12.00	12.00
He	10.99	10.93	10.98
C	8.56	8.52	8.46
N	8.05	7.92	7.90
O	8.93	8.83	8.76
Ne	8.09	8.08	7.95
Mg	7.58	7.58	7.62
Si	7.55	7.55	7.61
S	7.21	7.33	7.26
Ar	6.56	6.40	6.62
Ca	6.36	6.36	6.41
Fe	7.67	7.50	7.54
Ni	6.25	6.25	6.29

Galaxy is happening on a time scale very similar to that of Fe. They conclude that while in the early Universe the main C contributors are massive stars, C is later produced mainly by asymptotic giant branch (AGB) stars.

Nitrogen is produced during hydrogen burning via the CNO and CN cycles as both a primary and secondary element. In primary nucleosynthesis its production is independent of the initial metallicity of the star. Primary production of N happens during hydrogen shell burning in intermediate mass stars of ∼4–8 $M_\odot$ (Matteucci and Tosi 1985). Stellar models that include the effects of rotation indicate that massive stars between 9 and 20 $M_\odot$ may also produce primary N (Maeder and Meynet 2000). In secondary production, which is common to stars of all masses, N is synthesised from C and O and its abundance is therefore proportional to the initial metallicity of the star. The primary versus secondary production of N can be studied by investigating the N/O ratio as a function of O/H ratio. In the case of primary N production the N/O ratio will be constant. For the secondary production we will observe a linear correlation between the logarithms of N/O and O/H. The combination of primary and secondary nucleosynthesis will produce a non-linear relation. The ICM abundance of C and N was measured only in a few bright nearby galaxy clusters and elliptical galaxies (Peterson et al. 2003; Werner et al. 2006a). However, future instruments (see Sect. 8) will allow to measure the abundances of these elements in many clusters. This will be important for a better understanding of the stellar nucleosynthesis.

3 Abundance Studies Before *XMM-Newton* and *Chandra*

The discovery of the Fe-K line emission in the spectrum of the Perseus cluster by the *Ariel V* satellite (see the left panel of Fig. 3, Mitchell et al. 1976) and in Coma and Virgo by *OSO-8* (Serlemitsos et al. 1977) confirmed that the X-ray emission of galaxy clusters is predominantly thermal radiation from hot intra-cluster gas rather than inverse Compton radiation. These observations showed that the hot plasma in clusters of galaxies contains a significant portion of processed gas, which was ejected from stars in the cluster galaxies. Subsequent spectroscopic analysis of cluster samples observed with *OSO-8* and *HEAO-1 A2* satellites

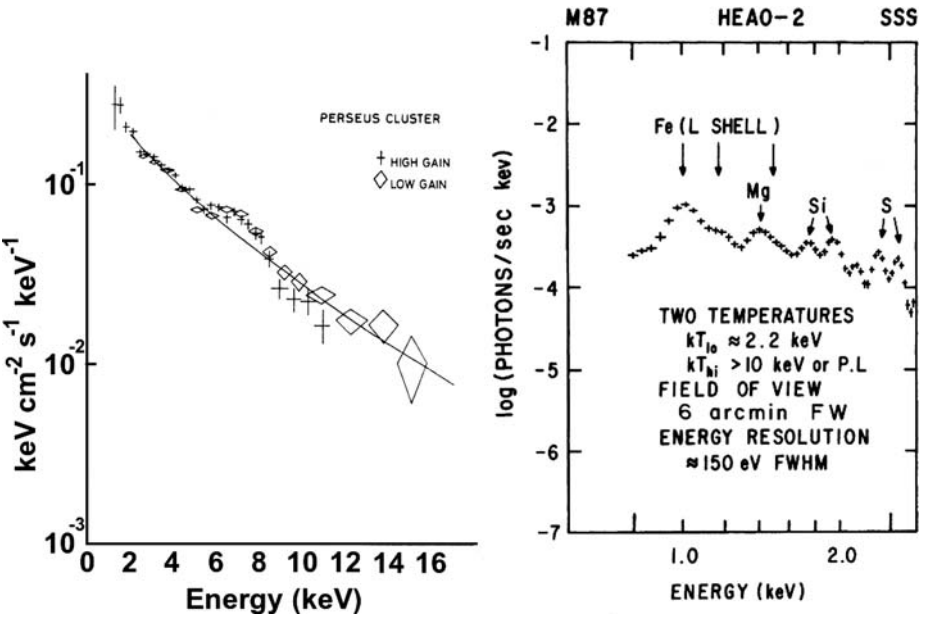


Fig. 3 *Left panel:* Spectrum of the Perseus cluster obtained by *Ariel 5*. The first cluster spectrum with observed Fe-K line emission (hump at ∼7 keV). From Mitchell et al. (1976). *Right panel:* Line emission observed in M 87 with the *Einstein* satellite. From Sarazin (1988), based on the work published by Lea et al. (1982)

revealed that the ICM has an Fe abundance of about one-third to one-half of the Solar value (Mushotzky et al. 1978; Mushotzky 1984). Combining spectra obtained by *Einstein* and *Ginga*, White et al. (1994) found an indication of a centrally enhanced metallicity in four cooling flow clusters. Unfortunately, the spectrometers on the *Einstein* observatory were not sensitive to the Fe-K line emission, because the mirror on the satellite was not sensitive to photon energies above ∼4 keV. However, the Solid-State Spectrometer (SSS) and the high-resolution Focal Plane Crystal Spectrometer (FPCS) on *Einstein* (Giacconi et al. 1979) allowed to detect emission lines at low energies. Using the SSS the K lines from Mg, Si, and S and the L lines from Fe were detected in the spectrum of M 87 (see the right panel of Fig. 3), Perseus, A 496, and A 576 (Lea et al. 1982; Mushotzky et al. 1981; Nulsen et al. 1982; Mushotzky 1984). Using the FPCS spectra of M 87 the O VIII Kα line was detected (Canizares et al. 1979), implying an O/Fe ratio of 3–5, and the relative strength of various Fe-L line blends showed that the gas cannot be at a single temperature.

However, until the launch of *ASCA* in 1993, Fe was the only element for which the abundance was accurately measured in a large number of clusters. *ASCA* allowed to detect the emission features from O, Ne, Mg, Si, S, Ar, Ca, Fe, and Ni in the spectra of a number of clusters. Furthermore, it allowed to accurately determine the Fe abundances out to redshift $z \approx 0.5$. *ASCA* data revealed a lack of evolution in the Fe abundance out to redshift $z \sim 0.4$ (Mushotzky and Loewenstein 1997; Rizza et al. 1998) and no evidence for a decrease at higher redshifts (Donahue et al. 1999). *ASCA* data were the first to show that in cooling core clusters the metallicity increases toward the centre (Fukazawa et al. 1994).

By analysing data of four clusters of galaxies, Mushotzky et al. (1996) found that the relative abundances of O, Ne, Si, S, and Fe are consistent with an origin in SN_{CC}. They suggest that SN_{CC} provided a significant fraction of metals in the ICM. Fukazawa et al.

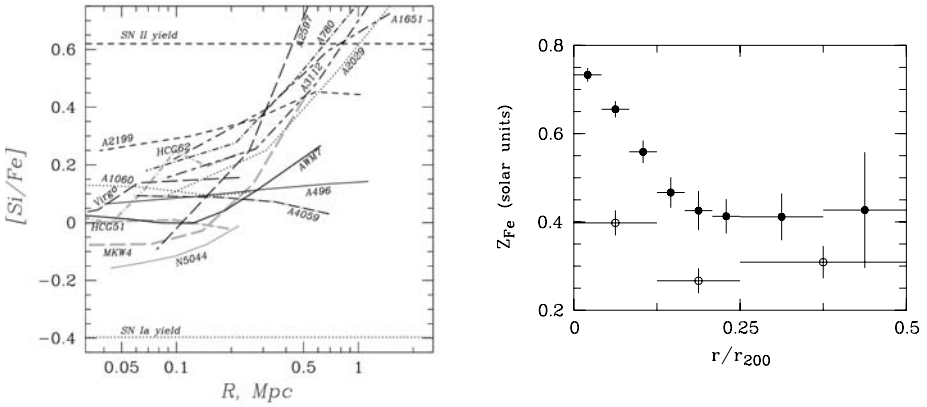


Fig. 4 *Left panel:* The radial distribution of the logarithm of the Si/Fe ratio expressed in Solar units (scale of Anders and Grevesse 1989) for a sample of clusters observed with *ASCA* (from Finoguenov et al. 2000). *Right panel:* The average radial distribution of the Fe abundance in Solar units (scale of Grevesse and Sauval 1998) for a sample of cooling core (*filled circles*) and non-cooling core (*empty circles*) clusters observed with *BeppoSAX* (from De Grandi et al. 2004)

(1998) showed that the Si abundance and the Si/Fe ratio increases from the poorer to the richer clusters, suggesting that the relative contribution of SN_{CC} increases toward richer clusters. They propose the possibility that a considerable fraction of SN_{CC} products escaped the poorer systems. Finoguenov et al. (2000) used *ASCA* data to show that SN Ia products dominate in the cluster centres and the SN_{CC} products are more evenly distributed (see the left panel of Fig. 4).

However, the large and energy dependent point-spread function ($\sim 2'$ half-power radius) of *ASCA* did not allow to investigate the spatial abundance distributions in detail. De Grandi and Molendi (2001) took advantage of the better spatial resolution ($\sim 1'$ half-power radius) of *BeppoSAX* and they measured the radial Fe abundance profiles for a sample of 17 rich nearby clusters of galaxies. They found that while the eight non-cooling core clusters in their sample have flat Fe abundance profiles, the Fe abundance is enhanced in the central regions of the cooling core clusters (see the right panel of Fig. 4). De Grandi et al. (2004) show that the mass associated with the abundance excess found at the centre of cool core clusters can be entirely produced by the brightest cluster galaxy (BCG).

The bandpass of the *ROSAT* PSPC was well suited for X-ray bright galaxies and groups of galaxies with temperatures of $\sim$1 keV and its PSF allowed for spatially resolved spectral analysis on a half-arcminute scale. Buote (2000a) investigated the deprojected abundance gradients in 10 galaxies and groups. In 9 of the 10 systems they found an abundance gradient, with the Fe abundance one to several times the Solar value within $\sim$10 kpc, decreasing to $\sim$0.5 Solar at 50–100 kpc distance.

The emerging picture from the abundance studies with *ASCA* and *BeppoSAX* data was that of an early homogeneous enrichment by SN_{CC}, which produce α-elements in the protocluster phase, and a subsequent more centrally peaked enrichment by SN Ia which have longer delay times and continue to explode in the cD galaxy for a long time after the cluster is formed.

Analysing and stacking all the data of clusters of galaxies observed with *ASCA* into temperature bins, Baumgartner et al. (2005) found a trend showing that the relative contribution of SN_{CC} is larger in higher temperature clusters than in the clusters with lower temperature.

Moreover, they found that Si and S do not track each other as a function of temperature, and Ca and Ar have much lower abundances than expected. They conclude that the α-elements do not behave homogeneously as a single group and the trends indicate that different enrichment mechanisms are important in clusters of different masses.

4 The Possibilities and Limitations of Elemental Abundance Determinations

4.1 Spectral Modelling and the Most Common Biases

The correct modelling of the temperature structure is crucial for the elemental-abundance determinations. Fitting the spectrum of a multi-temperature plasma with a simple single-temperature model results in best-fitting elemental-abundances that are too low. This effect is the most important in galaxies, groups of galaxies and in the cooling cores of clusters. Many studies have found significantly subsolar values for the Fe abundance in groups of galaxies, with abundances generally less than the stellar values in these systems, and less than those observed in galaxy clusters. To reconcile these observations with the chemical enrichment models, it has been postulated that galaxy groups accrete primordial gas after they have been expelling gas during most of their evolution (Renzini 1997). Buote and Canizares (1994) showed that if a galaxy spectrum has intrinsically two temperature components, the single-temperature model can give a significantly underestimated metallicity. Buote (2000b) demonstrated that in a spectrum characterised by two temperatures, one below 1 keV and one above 1 keV, with an average of $\sim$1 keV, the lower temperature component preferentially excites emission lines in the Fe-L complex below 1 keV (Fe XVII–XXI), while the higher temperature component excites the Fe-L lines from $\sim$1–1.4 keV (Fe XXI–XXIV). If these components contribute approximately equally to the emission measure, the spectral shape of the Fe-L complex will be broader and flatter than the narrower Fe-L complex produced by a single-temperature model with the average temperature of the two components. In order to fit the flat spectral shape at 1 keV, the single-temperature model will fit a lower Fe abundance and will increase the thermal continuum. The strongest bias is at the low temperatures, for higher temperature plasmas the bias is smaller as the Fe-K line-emission becomes more important for the determination of the Fe abundance. Buote (2000b) also shows an excess in the Si/Fe and S/Fe ratios inferred for two-temperature plasmas fit by single-temperature models. He notes that the lower temperature component in the two-temperature plasma produces stronger Si and S emission, and the single-temperature model underestimates the continuum at higher energies. Therefore the Si/Fe and S/Fe ratios of multi-temperature ICM inferred using single-temperature models will yield values in excess of the true values. Buote (2000b) shows that this effect is the most important for the hotter clusters of galaxies. This bias may have played a role in the excess Si/Fe ratios and trends derived using ASCA data (Mushotzky et al. 1996; Fukazawa et al. 1998; Finoguenov et al. 2000; Baumgartner et al. 2005).

Several authors (e.g., Buote et al. 2003; Werner et al. 2006b; Matsushita et al. 2007a), using XMM-Newton and Chandra data, showed that the best-fitting abundances in the multi-phase cooling core regions of clusters and in groups increase when two-temperature models or more complicated differential emission measure (DEM) models are used. Buote et al. (2003) show that the abundances of O, Mg, Si, and Fe in the centre of NGC 5044 obtained using a two-temperature model are higher than those obtained using a single-temperature model. They also show that the Si/Fe ratio in the centre is significantly lower when a two-temperature model is used. Similar results were obtained for NGC 1399 (see the left panel

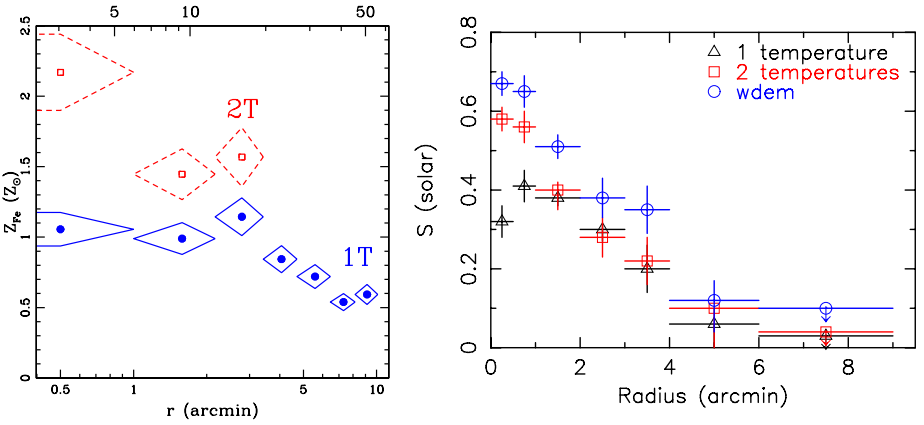


Fig. 5 These plots illustrate the influence of the temperature model on the best-fit abundances. *Left panel:* Radial Fe abundance distribution in NGC 1399 modelled with single- and two-temperature models (from Buote 2002). *Right panel:* Radial S abundance distribution in the cluster 2A 0335+096 modelled with single, two, and multi-temperature models (*triangles*, *squares*, and *circles*, respectively; from Werner et al. 2006b). Abundances in both figures are in the Solar units of Anders and Grevesse (1989)

of Fig. 5; Buote 2002). Analysing deep XMM-Newton data of the cluster 2A 0335+096, Werner et al. (2006b) showed that the abundances of all fitted elements (Mg, Si, S, Ar, Ca, Fe, Ni) in their model increased as they went from single- to two-temperature model, and to the multi-temperature DEM model. However, not only their absolute values increased but also the Si/Fe and the S/Fe ratios changed. Matsushita et al. (2007a) show an increase from single- to two- to multi-temperature models for O, Mg, Si, and Fe in the core of the Centaurus cluster. In the right panel of Fig. 5 we show the radial S profile obtained for 2A 0335+096 (Werner et al. 2006b) with single-, two-, and multi-temperature models.

As discussed in Buote (2000a) and Buote et al. (2003) in the lower mass systems ($\sim$ 1 keV) the best fit Fe abundance is sensitive to the lower energy limit used in spectral fitting. If the continuum at the low energies is not properly modelled, the equivalent widths of the lines in the Fe-L complex (and the oxygen line emission) will be inaccurate and the abundance determination will be incorrect. Buote et al. (2003) show the best-fitting Fe abundance in NGC 5044 determined with the lower energy limit at 0.7, 0.5, and 0.3 keV. They show that the results for each energy limit are significantly different and they obtain the highest Fe abundance for the lowest energy limit, which allows them to properly fit the continuum.

Since the $\sim$ 0.2 keV component of the Galactic foreground emits O VIII line emission, the uncertainties in the foreground/background model can cause systematic uncertainty in determining the oxygen abundance in the ICM. The temperature and flux of the $\sim$ 0.2 keV Galactic foreground component is position dependent on the sky and cannot be properly subtracted with blank fields (for detailed discussion see de Plaa et al. 2006).

Systematic uncertainties in the abundance determination can be introduced if bright point sources are not excluded from the spectral fits, especially in the lower surface brightness regions of clusters. The contribution of point sources raises the continuum level, systematically reducing the observed line equivalent widths. If their contribution is not properly modeled, the best fitting thermal model will have a higher temperature and lower elemental abundances.

In general two plasma models are used for fitting cluster spectra: MEKAL (Mewe et al. 1985, 1986; Kaastra 1992; Liedahl et al. 1995; Mewe et al. 1995) and APEC (Smith et al. 2001). These codes have differences in their line libraries and in the way of modelling the atomic physics. This allows to test the robustness of the abundance determinations. Systematic effects on the best fit abundances due to the adopted plasma model were investigated for example by Buote et al. (2003), Sanders and Fabian (2006), and de Plaa et al. (2007). They conclude that despite small differences in the determined temperature structure, the derived abundances are generally consistent within the errors.

Using simulated galaxy clusters Kapferer et al. (2007a) show that the more inhomogeneously the metals are distributed within the cluster, the less accurate is the observed metallicity as a measure for the true metal mass. They show that in general the true metal mass is underestimated by observations. If the distribution of metals in the ICM is inhomogeneous, the true metal mass in the central parts ($r < 500$ kpc) of galaxy clusters can be up to three times higher than the metal mass obtained by X-ray observations. The discrepancies are due to averaging. As the metallicity and temperature are not constant throughout the extraction area, the integration of thermal Bremsstrahlung and of line radiation can lead to underestimated metal masses.

4.2 Possibilities and Limitations of Current Instruments

XMM-Newton with its large effective area and superb spectral resolution, *Chandra* with its unprecedented spatial resolution, and *Suzaku* with its low background and good spectral resolution at low energies largely enhanced our knowledge on chemical abundances in galaxies, groups, and clusters of galaxies. In this section we describe the possibilities and the systematic uncertainties of abundance studies with these instruments.

4.2.1 XMM-Newton

Due to its large effective area and good spectral resolution, *XMM-Newton* is currently the best suited instrument to study abundances in the ICM. *XMM-Newton* has three identical telescopes and two sets of detectors on board. The first set of instruments is the European Photon Imaging Camera (EPIC), consisting of two metal oxide semiconductor—MOS1 and MOS2—CCD arrays collecting 50% of light from two telescopes, and the EPIC-pn CCD array collecting 100% of light from one telescope on board. The second set of instruments are the Reflection Grating Spectrometers (RGS) that collect 50% of light from two telescopes. Both sets of instruments play an important role in the abundance studies.

The advantages of EPIC are its broad bandpass, and good spatial and spectral resolutions, that allow us to detect the emission lines from O, Ne, Mg, Si, S, Ar, Ca, Fe, and Ni (see the top panel of Fig. 6). Furthermore, for deep observations of bright clusters, EPIC allows us to derive the spatial distribution of elements (see Sect. 5).

Unfortunately, the abundances determined in the cluster outskirts with low surface brightness using *XMM-Newton* remain uncertain due to the high background level. The value of the best fit temperature in these regions is very sensitive to the level of the subtracted background (e.g., de Plaa et al. 2006) and as we show above in Sect. 4.1, the temperature modelling is very important for the correct abundance determination.

The O line emission in the spectrum is strong, but its abundance is difficult to measure because it is sensitive to the assumed model for the Galactic foreground (see Sect. 4.1) and there are still remaining calibration uncertainties at low energies. The 1s–2p Ne lines at 1.02 keV are in the middle of the Fe-L complex (lying between 0.8 to 1.4 keV). The

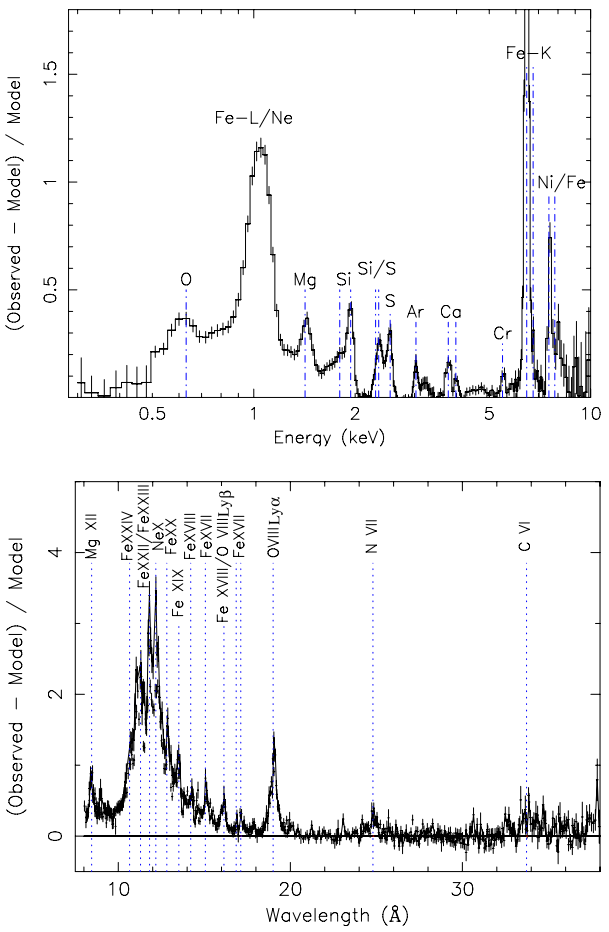


Fig. 6 *Top panel:* The line spectrum of the cluster 2A 0335+096, as observed with *XMM-Newton* EPIC (from Werner et al. 2006b). *Bottom panel:* Line spectrum of M 87, as observed with *XMM-Newton* RGS (from Werner et al. 2006a). After the fitting, the line emission was set to zero in the multi-temperature plasma model indicated on the y-axis. While the CCD spectra are plotted as a function of the observed energy, the grating spectra are shown as a function of the observed wavelength

resolution of the EPIC cameras is not sufficient to resolve the individual lines in the Fe-L complex, which makes the Ne abundance determination by EPIC uncertain. However, the abundances of O and Ne in the cores of cooling core clusters, observed with *XMM-Newton* with sufficient statistics, can be well determined using the high spectral resolution of the RGS.

The systematic differences in the effective-area between EPIC-MOS and EPIC-pn are of the order of 5–10% in certain bands (Kirsch 2006). By using data with high photon statistics and fitting spectra obtained by EPIC-MOS and EPIC-pn separately, de Plaa et al. (2007) investigated the magnitude of the systematic errors on the best-fitting abundances of elements heavier than Ne due to the uncertainties in the effective area calibration. They found that the elemental abundances most influenced by the calibration uncertainties are Mg, Si, and Ni. de Plaa et al. (2007) show that the systematic uncertainties on Mg are the largest, since the effective-area of EPIC-pn has a dip with respect to the effective-area of EPIC-MOS at the energies of the magnesium lines. This causes the Mg abundance derived by EPIC-pn to be underestimated. For the abundances of Si and Ni they estimate that the systematic uncertainties are 11% and 19%, respectively. Although EPIC-MOS and EPIC-pn show large systematic differences at the energy of the Mg lines, the Mg abundances derived by EPIC-

MOS and *Chandra* ACIS for the Centaurus cluster are consistent (Matsushita et al. 2007a; Sanders and Fabian 2006), and the Mg abundances derived by the EPIC-MOS and RGS using the deep observations of 2A 0335+096 also agree with each other.

The S abundance determination with EPIC is robust (e.g., de Plaa et al. 2007). Ar and Ca abundances should also be accurately determined, however their signal is usually weak, so the level of their observed equivalent widths may be somewhat sensitive to the continuum level.

The RGS has a high spectral resolution but only a limited spatial resolution in the cross-dispersion direction. It allows for relatively accurate abundance measurements of O, Ne, Mg, Fe, and in the case of deep observations of nearby bright clusters and elliptical galaxies even the spectral lines of C and N are detected (see the bottom panel of Fig. 6). In the case of deep observations of nearby bright clusters, RGS allows us to extract spectra from typically five bins that are $\sim 1'$ wide in the cross-dispersion direction and the photons in each bin are collected from $\sim 10'$ long regions in the dispersion direction. This way RGS allows us to determine projected abundance profiles.

The spectral resolution of the RGS is, however, limited by the broadening of the observed line profiles caused by the spatial extent of the source along the dispersion direction. This makes RGS spectroscopy useful only for cooling core clusters with strongly peaked surface brightness distribution. In order to account for the line broadening in the spectral modelling, the model is convolved with the surface brightness profile of the source along the dispersion direction. Because the radial profile for a spectral line produced by an ion can be different from the radial surface brightness profile in a broad band, in practice the line profile is multiplied with a scale factor, which is a free parameter in the spectral fit. Different best fit line profiles for the Fe lines and for the O line emission can suggest a different radial distribution for these elements (de Plaa et al. 2006).

4.2.2 Chandra

Although the Advanced CCD Imaging Spectrometer (ACIS) on *Chandra* has a much smaller effective area than the EPIC on *XMM-Newton*, which makes it less suitable for abundance studies of the ICM, it is the instrument of choice for studies of metal enrichment in galaxies, where a good spatial resolution is required. *Chandra* is able to resolve the substantial fraction of X-ray binaries in galaxies as point sources, so their contribution can be excluded from the spectral modelling of the inter-galactic medium (IGM). The excellent spatial resolution also makes it possible to investigate the spatial temperature variations, which can bring a bias into the determination of abundances (see Sect. 4.1). The spectral resolution of ACIS is sufficient to reliably determine the abundances of a number of elements in the ICM and IGM. With deep enough exposure times *Chandra* ACIS can be used to study the ICM abundances of O, Ne, Mg, Si, S, Ar, Ca, Fe, and Ni (e.g., Sanders et al. 2004; Sanders and Fabian 2006).

Humphrey et al. (2004) discuss the systematic uncertainties in the abundance determinations due to the uncertainties in the absolute calibration of ACIS. After correcting for the charge transfer inefficiency (CTI) using an algorithm by Townsley et al. (2002), they found a change of $\Delta Z_{\rm Fe} \simeq 0.3$ in the best fit Fe abundance of NGC 1332.

Sanders and Fabian (2006) investigate the systematic uncertainties on the abundance determinations with the S3 chip of *Chandra* ACIS in the Centaurus cluster. They conclude that for the best determined elements the non-CTI-corrected *Chandra* and *XMM-Newton* results agree well. The only element where the *Chandra* and *XMM-Newton* data do not agree is oxygen. There has been a buildup of contamination on the ACIS detectors, which influences

the effective area calibration at the low energies making the O abundance determination difficult. The iridium coating of the *Chandra* mirrors produces a steep drop in the effective area at around 2.05 keV, introducing a potential systematic uncertainty in the Si abundance determination. However, the Si abundances determined by *Chandra* and *XMM-Newton* agree well (Sanders and Fabian 2006).

4.2.3 Suzaku

The advantage of the X-ray Imaging Spectrometer (XIS) detectors (Koyama et al. 2007) on *Suzaku* (Mitsuda et al. 2007) compared to EPIC on *XMM-Newton* is its better spectral resolution below 1 keV, and the low background, providing a better sensitivity at the O line energy. *Suzaku* also allows for better measurements of the Mg lines in the low surface brightness cluster outskirts, which is difficult with EPIC/MOS due to the presence of a strong instrumental Al line close to the expected energy of the observed Mg lines.

The relatively broad point spread function of XIS with a half-power diameter of $\sim 2'$ (Serlemitsos et al. 2007) does not allow detailed investigations of the spatial distribution of elements.

The effective area of the XIS detectors below 1 keV is affected by carbon and oxygen contamination of the optical blocking filter. The contamination is time and position dependent. Its column density and spatial distribution are still not well known. Matsushita et al. (2007b) estimate that the uncertainty in the thickness and chemical composition of the contaminating column results in a $\sim$20% systematic error on the O abundance.

5 Spatial Distribution of Elements

5.1 Radial Abundance Profiles

Observations with *XMM-Newton* and *Chandra* confirmed the centrally peaked metallicity distribution in the cooling core clusters and the flat distribution of metals in the non-cooling core clusters (Vikhlinin et al. 2005; Pratt et al. 2007) previously seen with *BeppoSAX*. *XMM-Newton* and *Chandra* also confirmed the strong contribution of SN Ia to the enrichment of cluster cores found by *ASCA*. The observations are consistent with the picture of an early contribution by SN_{CC} and later enrichment of cluster cores by SN Ia. One of the first abundance studies with *XMM-Newton*, performed on the cooling core cluster Abell 496, showed a radially flat distribution of O, Ne, and Mg and centrally peaked distribution of Ar, Ca, Fe, and Ni (Tamura et al. 2001). The abundance peak in the cluster core is consistent with the idea that the excess metals were produced by SNe Ia in the cD galaxy. However, the strongly increasing Si/Fe ratio toward the outskirts seen with *ASCA* was not confirmed with *XMM-Newton*.

Analysing *XMM-Newton* EPIC data of M 87, Böhringer et al. (2001) and Finoguenov et al. (2002) found rather low O/Fe ratio in the cluster core, that increased toward the outer regions. Analysing the same dataset, Gastaldello and Molendi (2002) questioned these results, showing that the O/Fe ratio in the inner $9'$ is constant. As shown in the left panel of Fig. 7, using *XMM-Newton* RGS data, Werner et al. (2006a) found a centrally peaked Fe abundance gradient and a flat O distribution in M 87 confirming the results of Böhringer et al. (2001), Finoguenov et al. (2002), and Matsushita et al. (2003) (the O abundance profiles found by Gastaldello and Molendi (2002) were different because of calibration problems at low energies in the early days of the mission). The RGS data reveal an O/Fe ratio

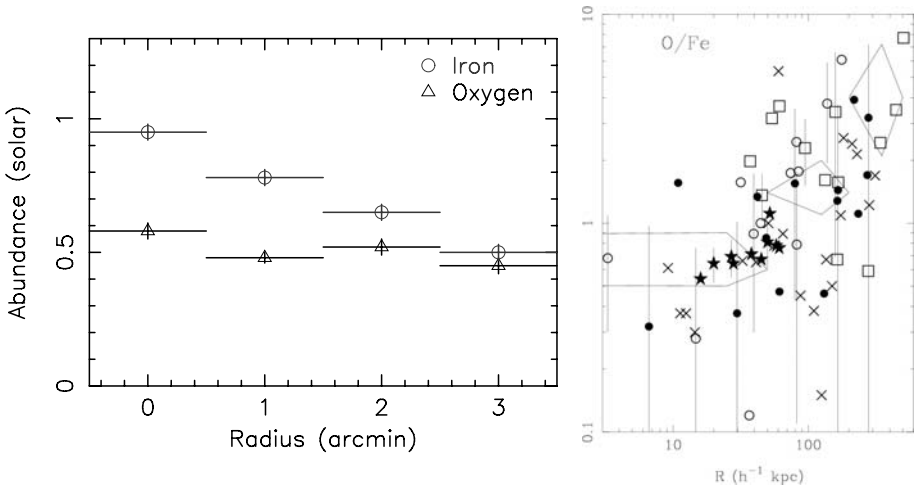


Fig. 7 *Left panel:* Radial distribution of Fe and O in M 87 determined using *XMM-Newton* RGS (from Werner et al. 2006a). The abundances are shown relative to the proto-solar values given by Lodders (2003). *Right panel:* Radial distribution of the O/Fe ratio determined for a sample of clusters using *XMM-Newton* EPIC (from Tamura et al. 2004). The abundance ratios are shown relative to the Solar values given by Anders and Grevesse (1989)

of 0.6 in the cluster core and a ratio of 0.9 at a distance of 3′ (14 kpc) from the centre. However, the *XMM-Newton* EPIC data show that the Si abundance has a similar gradient in M 87 as the Fe abundance. A similar abundance pattern with a rather flat radial distribution of O, and centrally peaked Si, and Fe abundances is observed in the cluster sample analysed by Tamura et al. (2004 see the right panel of Fig. 7), in the Perseus cluster (Sanders et al. 2004), in Abell 85 (Durret et al. 2005), in Sérsic 159-03 (de Plaa et al. 2006), in 2A 0335+096 (Werner et al. 2006b), in the Centaurus cluster (Matsushita et al. 2007a; Sanders and Fabian 2006), and in the Fornax cluster (Matsushita et al. 2007a). The abundances of Si and Fe have similar gradients also in the group of galaxies NGC 5044 (Buote et al. 2003), with a constant radial distribution of Si/Fe = 0.83. The radial distribution of O in this group is consistent with being flat. *Suzaku* observations of the clusters A 1060 (Sato et al. 2007c) and AWM 7 (Sato et al. 2007a) also showed rather low O abundance of around 0.5 Solar within about 5′ from the core without a clear abundance decline in the outer regions.

The radial metallicity gradients in cooling flow clusters often display an inversion in the centre (e.g., Sanders and Fabian 2002). This apparent metallicity drop in the very central region is often the result of an oversimplified model in the spectral analysis (see Sect. 4 and Molendi and Gastaldello 2001). However, in some cases the metallicity drop in the cluster core is robust. In the Perseus cluster the abundances peak at the radius of 40 kpc and they decrease at smaller radii. This inversion does not disappear when extra temperature components and power-law models are included, or when projection effects are taken into account (Sanders and Fabian 2007).

Böhringer et al. (2004) found that long enrichment times ($\gtrsim$5 Gyr) are necessary to produce the observed central abundance peaks. Since the classical cooling flows, or a strongly convective intra-cluster medium, or a complete metal mixing by cluster mergers would destroy the observed abundance gradients, they conclude that cooling cores in clusters are preserved over very long times. Böhringer et al. (2004) and Matsushita et al. (2007b) show that the iron-mass-to-light ratio in the central region of the Centaurus cluster (the mass of Fe

relative to the B-band luminosity of the stellar population of the cD galaxy) is much higher than that of M 87. This indicates that the accumulation time scale of the SN Ia products was much higher in Centaurus than in M 87. In agreement with the conclusions of De Grandi et al. (2004) they also conclude that the innermost part of the ICM is dominated by gas originating mainly from the stellar mass loss of the cD galaxy.

If the metals originate from the stellar population of the cD galaxy, then in the absence of mixing the abundance profiles would follow the light profiles. The observed metal profiles are much less steep than the light profiles, which suggests that the injected metals are mixed and they diffuse to larger radii. Rebusco et al. (2005, 2006) model the diffusion of metals by stochastic gas motions after they were ejected by the brightest galaxy and compare their results with the peaked Fe abundance profiles of 8 groups and clusters. They suggest that the AGN/ICM interaction makes an important contribution to the gas motion in the cluster cores.

In order to determine the metallicities of the stellar populations of galaxies and star clusters it has become a common practice to use the feature made up of the Mg b line and the Mg H band around 5200 Å, which is known as the Mg_2 index. The Mg_2 index appears to be a relatively good metallicity indicator. In M 87, Centaurus, and the Fornax clusters the O and Mg abundances in the centre are consistent with the optical metallicity of the cD galaxy derived from the Mg_2 index (Matsushita et al. 2003, 2007a, 2007b). The Mg_2 index mainly depends on the Mg abundance, but also on the total metallicity where the O contribution matters most, and depends weakly on the age population of stars. The agreement between the O and Mg abundances in the ICM with the stellar metallicities supports the view that the ICM in the cluster centres is dominated by the gas ejected from the cD galaxies. However, the Mg/Fe ratio in the stellar populations of elliptical galaxies is significantly higher than that in the hot gas (Humphrey and Buote 2006). The observed difference shows that since the time when the current stellar populations of these ellipticals formed, a significant number of SNe Ia enriched the gas in addition to the stellar mass loss.

Observations indicate that at large radii clusters have a flat abundance distribution of about 0.2 Solar. By using *Suzaku* data, Fujita et al. (2007) found that the gas is enriched to a metallicity of ∼0.2 Solar out to the virial radii of clusters A 399 and A 401. These clusters are at the initial stage of a merger and their virial radii partially overlap. The authors interpret their finding as a result of early enrichment by galactic superwinds in the proto-cluster stage. However, the extraction region used by Fujita et al. (2007) is large and their measured Fe abundance might be significantly biased toward the value at about half of the virial radii of the clusters.

5.2 2D Distribution of Metals

Deep observations of bright clusters of galaxies with *XMM-Newton* and *Chandra* allow to map the 2D distribution of metals in the ICM (Sanders et al. 2004; Durret et al. 2005; O'Sullivan et al. 2005; Sauvageot et al. 2005; Finoguenov et al. 2006; Werner et al. 2006b; Sanders and Fabian 2006; Bagchi et al. 2006). In general the distribution of metals in clusters is not spherically symmetric, and in several cases it shows edges in the abundance distribution or several maxima and complex metal patterns. The 2D distribution of metals is a good diagnostic tool, which can be used together with the maps of thermodynamic properties to investigate the complex structure of cluster cores and the merging history of clusters (e.g., Kapferer et al. 2006).

M 87 provides us with the unique opportunity to study in detail the role of the AGN feedback in transporting and distributing the metals into the ICM. The high X-ray surface

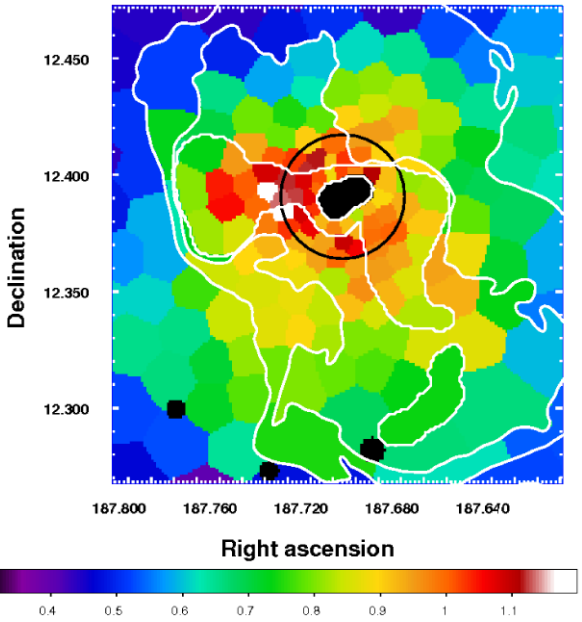


Fig. 8 Map of the Fe abundance in M 87. The *colour bar* indicates the Fe abundance with respect to the proto-solar value of Lodders (2003), with *red* showing higher and *blue* lower abundances. Contours of the 90 cm radio emission (Owen et al. 2000) are overplotted. The *black circle* indicates the half light radius of M 87. Beyond the expected radial gradient, one clearly sees the enhanced Fe abundance in the radio arms, especially within the Eastern arm. From Simionescu et al. (2007)

brightness and small distance of M 87 allow a detailed study of the 2D distribution of several elements (O, Si, S, and Fe). By analysing deep (120 ks) XMM-Newton data, Simionescu et al. (2007) found that the spatial distribution of metals shows a clear enhancement along the radio lobes, where the rising radio emitting plasma ejected by the AGN uplifts cooler gas from the core of the cluster (see Fig. 8). Furthermore, they showed that the metallicity of the cool uplifted ($kT < 1.5$ keV) gas is ∼2.2 times Solar and the abundance ratios in and outside the multi-phase region associated with the uplifted gas are very similar. They estimate that the mass of the cool gas is $\approx 5 \times 10^8\ M_\odot$ and it probably originates from the stellar winds enriched with SN Ia products. Approximately 40–140 Myr are required to produce the metals seen in the cool gas, indicating that the uplift of cool gas by AGN radio bubbles is a rare event.

6 Element Ratios and Their Reconstruction with Supernova Models

The chemical elements that we see in the ICM are the integral yield of all the different supernova types that have left their specific abundance patterns in the gas prior and during cluster evolution. Since the launch of *ASCA*, which for the first time allowed the determination of abundances of elements other than Fe in the ICM (see Sect. 3), several groups tried to use the ICM abundance patterns to constrain the contribution of different kinds of supernovae to the ICM enrichment and to put constraints on the theoretical supernova models.

In the following subsection we review these efforts. The still somewhat ambiguous results of these efforts are summarised in Sect. 6.2

6.1 Results on Reconstructed Supernova Models

Mushotzky et al. (1996) investigated the elemental abundances in four clusters, removing the central regions, and concluded that the abundance ratios are consistent with the contribution

of SN_{CC}. Dupke and White (2000) found that most of the Fe (∼60–70%) in the cores of their investigated clusters comes from SN Ia, and they used these central regions dominated by SN Ia ejecta to discriminate between competing theoretical models for SN Ia. They showed that the observed high Ni/Fe ratio of ∼4 is more consistent with the W7 convective deflagration SN Ia model than with the delayed detonation models. Stacking all the cluster data in the *ASCA* archive into temperature bins, Baumgartner et al. (2005) showed that the abundances of the most well determined elements Fe, Si, and S do not give a consistent solution for the fraction of material produced by SN Ia and SN_{CC} at any given cluster temperature. They concluded that the pattern of elemental abundances requires an additional source of metals other than SN Ia and SN_{CC}, which could be the early generation of Population III stars.

The *ASCA* estimates were largely based on the Si/Fe ratio, which has the disadvantage of Si being contributed by both SN Ia and SN_{CC}. The much improved sensitivity of *XMM-Newton* to measure the spectral lines of more elements at a higher statistical significance offers better opportunities to constrain the supernova models.

Finoguenov et al. (2002) used *XMM-Newton* EPIC to accurately measure the abundances of O, Ne, Mg, Si, S, Ar, Ca, Fe, and Ni in the hot gas in two radial bins of M 87. They found that the Si-group elements (Si, S, Ar, Ca) and Fe have a stronger peak in the central region than O, Ne, and Mg. The SN Ia enrichment in the centre is characterised by a Solar ratio of Si-group elements to Fe. The inferred SN Ia enrichment in the outer region, associated with the ICM of the Virgo cluster, has a lower ratio of Si-group elements to Fe by 0.2 dex, which is more characteristic for the ICM of other clusters (Finoguenov et al. 2000). Finoguenov et al. (2002) also find a Ni/Fe ratio of ∼1.5 in the central region, while they point out that this ratio for other clusters determined by *ASCA* is ∼3. They conclude that these abundance patterns confirm the diversity of SN Ia and to explain the SN Ia metal enrichment in clusters both deflagration and delayed detonation scenarios are required. The high SN Ia yield of Si-group elements in the centre may imply that the Si burning in SN Ia is incomplete. This favours the delayed-detonation models with lower density of deflagration-detonation transition, lower C/O ratio, and lower central ignition density. On the other hand, the abundance patterns in the outer region of M 87 are more characteristic of deflagration supernova models.

A bimodal distribution is observed in the magnitude decline of the SN Ia events. While a population of SN Ia with a slow decline is more common in spiral and irregular galaxies with recent star formation, a fainter and more rapidly decaying population of SN Ia is more common in early-type galaxies (Hamuy et al. 1996; Ivanov et al. 2000). The SN Ia with incomplete Si burning enriching the central part of M 87 could be associated with optically fainter SNe Ia observed in early-type galaxies. Mannucci et al. (2006) found that the present data on SN Ia rates in different populations of parent galaxies indicate that SN Ia have a bimodal delay time distribution. They conclude that about 50% of SN Ia explode soon after the formation of the progenitor binary system, in a time of the order of 10^8 years, while the delay time distribution of the remaining 50% of SN Ia can be described by an exponential function with a decay time of ∼3 Gyr. The supernovae with short delay times might be associated with the supernovae with the slower decline, while the supernovae with long delay times might be associated with the rapidly decaying SNe Ia. If this association is true, then the observed Si/Fe of the accumulated SN Ia ejecta will be a function of the age of the system.

Matsushita et al. (2007a) show that SN Ia ejecta have a higher Fe/Si ratio in regions with larger iron-mass-to-light ratio (IMLR). Higher IMLR corresponds to longer accumulation time scale. The correlation between the inferred Fe/Si ratio of SN Ia ejecta and the IMLR

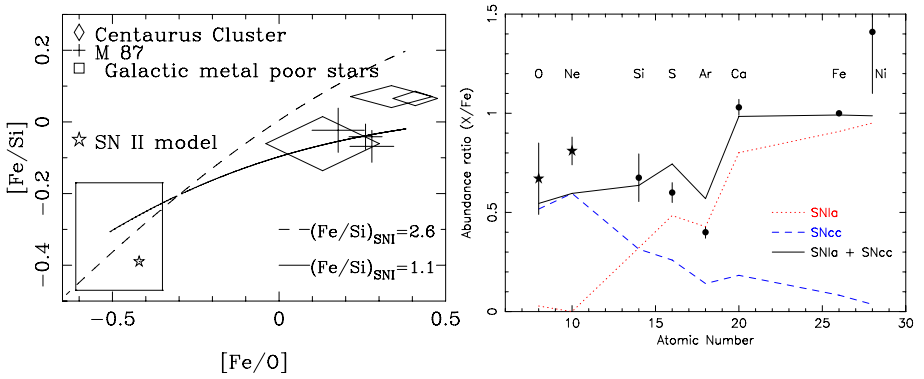


Fig. 9 *Left panel:* Comparison of the [Fe/Si] ratio of the ICM in the Centaurus cluster and in M 87, plotted against [Fe/O], from Matsushita et al. (2007a). Abundances are here on the scale of Feldman (1992). The *open square* shows the average value for Galactic metal poor stars (Clementini et al. 1999) and the *asterisk* shows the abundance ratios for SN_{CC} calculated by Nomoto et al. (1997). The *solid line* and *dashed line* show the abundance pattern synthesised by SN Ia with Fe/Si = 1.1 (best fit relation for the core of M 87) and Fe/Si = 2.6 (the ratio produced by the W7 model Nomoto et al. 1984), respectively. *Right panel:* Reconstruction of the abundance patterns (at the proto-solar Lodders 2003 scale) observed in the sample of 22 clusters of de Plaa et al. (2007) with theoretical supernova yields. The abundance values indicated by *filled circles* are obtained with *XMM-Newton* EPIC. The abundances of O and Ne are the average values obtained for Sérsic 159-03 and 2A 0335+096 with RGS. The SN Ia yields are for a delayed detonation model calibrated on the Tycho supernova remnant by Badenes et al. (2006). The SN_{CC} yields are for progenitors with Solar metallicities by Nomoto et al. (2006) integrated over the Salpeter IMF. From de Plaa et al. (2007)

supports the suggestion by Finoguenov et al. (2002) and Matsushita et al. (2003) that the average Fe/Si ratio produced by SN Ia depends on the age of the system.

Böhringer et al. (2005) compiled the available observational results of the O/Si/Fe from *XMM-Newton*, *Chandra*, and *ASCA* to show that the observed abundance patterns in clusters are more consistent with the WDD models, which provide larger Si/Fe ratios than the W7 model. Buote et al. (2003) and Humphrey and Buote (2006) also discuss the higher α-element enrichment observed in groups and in elliptical galaxies. They conclude that there is an increasing evidence that the O is over-predicted by the theoretical SN_{CC} yields, or there is a source of α-element enrichment in addition to that provided by SN_{CC} and SN Ia.

Using deep *XMM-Newton* observations of the clusters 2A 0335+096 and Sérsic 159-03, Werner et al. (2006b) and de Plaa et al. (2006) determined the abundances of 9 elements and fitted them with nucleosynthesis models for supernovae. They found that ∼30% of all supernovae enriching the ICM were SN Ia and ∼70% were SN_{CC}. They also found that the Ca abundance in these clusters is higher than that expected from the models. Later, de Plaa et al. (2007) used a sample of 22 clusters observed with *XMM-Newton* to determine the abundances of Si, S, Ar, Ca, Fe, and Ni in the ICM within the radius of $0.2R_{500}$ and they compared the best-fitting abundances with theoretically predicted yields of different SN Ia and SN_{CC} models (see Fig. 9). Because the SN_{CC} do not have a significant impact on the Ar/Ca ratio, de Plaa et al. (2007) used the Ar and Ca abundances in clusters as a criterion for determining the quality of SN Ia models. They found that the Ar and Ca abundances in the ICM are inconsistent with currently favoured SN Ia models. However, they showed that their best fit Ca/Fe and Ar/Ca ratios are consistent with the empirically modified delayed-detonation SN Ia model, which is calibrated on the Tycho supernova remnant (Badenes et al. 2006). Using this model, de Plaa et al. (2007) obtained a number ratio of

$N_{\mathrm{SN\,Ia}}/N_{\mathrm{SN\,Ia+SN_{CC}}} = 0.44 \pm 0.10$, which suggests that binary systems in the appropriate mass range are very efficient (∼5–16%) in eventually forming SN Ia explosions.

Based on *Suzaku* observations of 2 clusters (A1060 and AWM 7) and 2 groups of galaxies (HCG 62 and NGC 507), Sato et al. (2007b) obtained abundance patterns for O, Mg, Si, S, and Fe in the region out to $0.3r_{180}$. Contrary to most of the previously mentioned results, they found a better fit to the observed abundance patterns using the W7 SN Ia model rather than the WDD models. Assuming the W7 model for SN Ia, the yields from Nomoto et al. (2006) and the Salpeter IMF for $\mathrm{SN_{CC}}$, the ratio of occurrence numbers of $\mathrm{SN_{cc}}$ to SN Ia is ∼3.5. It corresponds to the number ratio of $N_{\mathrm{SN\,Ia}}/N_{\mathrm{SN\,Ia+SN_{CC}}} = 0.22$, which is smaller than that obtained by de Plaa et al. (2007). The number ratios were determined only considering the metals in the ICM and metals locked up in the stars were not considered.

6.2 Summary of the Efforts of Using ICM Abundances to Constrain Supernova Models

The abundance patterns obtained during many deep observations of clusters of galaxies with XMM-Newton clearly favour the delayed-detonation SN Ia models (see Böhringer et al. 2005) and the observed abundance ratios of Ar/Ca are relatively well reproduced using empirically modified delayed-detonation models calibrated on the Tycho supernova remnant (de Plaa et al. 2007). However, several observations still favour the W7 SN Ia model (Sato et al. 2007b) and in most of the cases the abundances are not determined well enough to really discriminate between the different models. As Finoguenov et al. (2002) suggest it might even be possible that both the deflagration and delayed-detonation types of SN Ia play a role in the chemical enrichment of clusters of galaxies.

XMM-Newton RGS observations of the hot gas in clusters of galaxies and in elliptical galaxies reveal Ne/O ratios which are higher than Solar (Xu et al. 2002; Peterson et al. 2003; Werner et al. 2006a, 2006b). SN Ia produce very little O and Ne, and their ratio can be used to put constraints on the progenitors of $\mathrm{SN_{CC}}$. From the $\mathrm{SN_{CC}}$ models discussed in Sect. 2 only the models with pre-enriched progenitors predict higher than Solar Ne/O ratio.

While it is well known that elements from O up to the Fe-group are primarily produced in supernovae, the main sources of C and N are still being debated (see Sect. 2). Deep observations of nearby X-ray bright elliptical galaxies and bright clusters with the RGS on *XMM-Newton* allow to measure the C and N abundances in the hot gas. By combining the RGS data obtained in two deep *XMM-Newton* observations of M 87, Werner et al. (2006a) determined relatively accurate C and N abundance values. The small O/Fe ratio and large C/Fe and N/Fe ratios found in M 87 suggest that the main sources of C and N are not the massive stars that also produce large quantities of O, but the low- and intermediate-mass asymptotic giant branch stars.

7 Abundances as Function of Cluster Mass and Redshift

Arnaud et al. (1992) studied the correlations between the gas mass, Fe mass, and optical luminosity in clusters of galaxies. They found that the gas mass in the ICM is highly correlated and the Fe mass is directly proportional to the integrated optical luminosity of elliptical and lenticular galaxies. They found no correlation with the integrated optical luminosity of spiral galaxies. The ratio of the gas mass in the ICM to the stellar mass increases with the cluster richness.

Renzini (1997) showed that the IMLR is very similar in all clusters of galaxies with ICM temperature above ∼1 keV. However, below ∼1 keV the IMLR seems to drop by almost 3

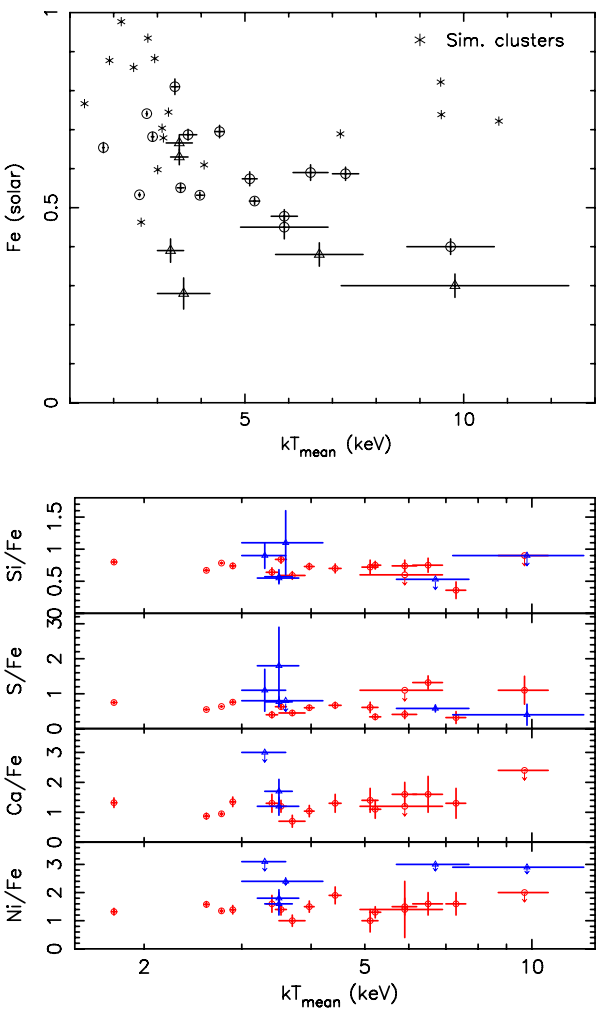


Fig. 10 *Top panel:* Best fit Fe abundance as a function of ICM temperature determined within $0.2R_{500}$ for a sample of clusters investigated by de Plaa et al. (2007). We compare the observed Fe abundances with values from simulations (Borgani et al. 2008—Chap. 18, this issue) for a Salpeter IMF, extracted within the same physical radius. *Bottom panel:* Abundance ratios of Si, S, Ca, and Ni with respect to Fe as a function of the cluster temperature, compiled from de Plaa et al. (2007). The abundance ratios are consistent with being constant as a function of cluster mass. Cooling core clusters are shown as *red circles* and non-cooling core clusters as *blue triangles*. Abundances in both panels are plotted with respect to the proto-solar abundance values of Lodders (2003)

orders of magnitude. This drop at low temperatures is due to the combination of lower gas mass and lower measured Fe abundance in cool systems. However, it was shown that the low Fe abundances measured in many groups were artifacts of wrong temperature modelling (see later).

Renzini (1997, 2004) showed that clusters hotter than ∼2.5 keV have similar abundances, with very small dispersion. Below 2.5 keV, the best fit metallicities show a large range of values. The previously reported very low abundance values for low temperature groups and elliptical galaxies turn out to be much higher when two- or multi-temperature fitting is applied (see Sect. 4.1 and Buote and Fabian 1998; Buote 2000b; Buote et al. 2003; Humphrey and Buote 2006). However, the outer regions of galaxy groups still show very low metallicities. An XMM-Newton observation of the group NGC 5044 indicates an iron abundance in the hot gas of $Z_{\rm Fe} \sim 0.15$ Solar in the region between 0.2 and 0.4 of the virial radius (Buote et al. 2004). This result shows that the total iron mass to optical light ratio in NGC 5044 is about 3 times lower than that observed in rich clusters.

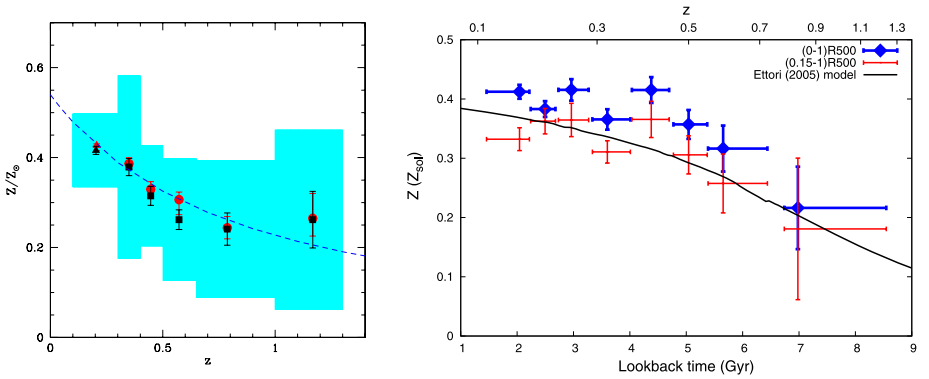


Fig. 11 *Left panel:* Mean Fe abundance from combined fits within six redshift bins (*red circles*) and the weighted average of single source observations in the same bins (*black square*) from Balestra et al. (2007). The shaded areas show the *rms* dispersion. The *dashed line* shows the best power-law fit over the redshift bins. *Right panel:* Mean ICM metal abundances in eight redshift bins within R_{500} with and without the central $0.15R_{500}$ excluded (from Maughan et al. 2008). The *solid line* shows the supernova enrichment model of Ettori (2005). Both panels use the abundance scale of Anders and Grevesse (1989)

We compiled the abundance values measured within $0.2R_{500}$ for the sample of 22 clusters analysed by de Plaa et al. (2007). In the top panel of Fig. 10 we show the best fit Fe abundance as a function of ICM temperature. Our data confirm earlier results showing that while in cooler clusters, in the temperature range of 2–4 keV, $Z_{\rm Fe}$ shows a range of values between 0.2–0.9 Solar, in hot massive clusters ($kT \gtrsim 5$ keV) the Fe abundance is on average lower and equal to $Z_{\rm Fe} \sim 0.3$ Solar. We compare the observed trend with values from simulations extracted within $0.2R_{500}$ as the *XMM-Newton* data (Borgani et al. 2008—Chap. 18, this issue). The data are simulated for the Salpeter IMF. The data are more consistent with the simulated values at the lower temperatures. At higher temperatures the simulated abundances are higher than the observed values. The observed trend might be linked to a decrease of the star formation efficiency with increasing cluster mass (Lin et al. 2003).

In the bottom panel of Fig. 10 we show the abundance ratios of Si, S, Ca, and Ni with respect to Fe as a function of the cluster temperature, compiled from de Plaa et al. (2007). The abundance ratios are consistent with being constant as a function of temperature, which is related to the cluster mass. The intrinsic spread in abundance ratios between clusters is smaller than 30%. This finding contradicts the *ASCA* results (Baumgartner et al. 2005), which show an increase of the Si/Fe ratio from 0.7 to 3 Solar in the 2–8 keV temperature range. This result indicates that the ratios of SN Ia and $\rm SN_{CC}$ contributing to the enrichment of the ICM are similar for clusters of all masses.

Investigating a sample of 56 clusters at $z \gtrsim 0.3$, Balestra et al. (2007) confirmed the trend of the Fe abundance with cluster temperature also in the higher redshift clusters. They show that the abundance measured within $(0.15–0.3)R_{\rm vir}$ in clusters below 5 keV is, on average, a factor of $\sim$2 larger than in hot clusters, following a relation $Z(T) \simeq 0.88T^{-0.47}$.

Balestra et al. (2007) also found an evolution of the Fe content in the ICM with redshift (see the left panel of Fig. 11). They conclude that the average Fe content of the ICM at the present epoch is a factor of $\sim$2 larger than at $z \simeq 1.2$. While at the redshift $z \gtrsim 0.5$ they observe an average ICM abundance of $Z_{\rm Fe} \simeq 0.25\ {\rm Z}_{\odot}$, in the redshift range of $z \simeq 0.3$–0.5 they measure $Z_{\rm Fe} \simeq 0.4\ {\rm Z}_{\odot}$. They parametrise the decrease of metallicity with redshift with a power-law $\sim (1+z)^{-1.25}$. The evolution in the Fe abundance with redshift was confirmed by

Maughan et al. (2008), who investigated a sample of 116 clusters of galaxies at $0.1 < z < 1.3$ in the *Chandra* archive (see the right panel of Fig. 11). They found a significant evolution with the abundance dropping by 50% between $z \sim 0.1$ and $z \sim 1$. They verified that the evolution is still present, but less significant, when the cluster cores are excluded from the abundance measurement (they use apertures of 0–1 R_{500} and 0.15–1 R_{500}), indicating that the evolution is not solely due to disappearance of relaxed, cool core clusters at $z \gtrsim 0.5$.

Simulations by Kapferer et al. (2007b) show that between redshifts of $z = 1$ and $z = 0$ a galaxy cluster accretes fresh gas, with the gas mass increasing during this time period by a factor of ≈3. Which means that in order to get an increase of the metallicity during this time period the freshly accreted gas must be pre-enriched or the input of metals into the ICM in this time interval must be larger than expected.

8 Future of ICM Abundance Studies

A major improvement in X-ray spectroscopy will be the employment of micro-calorimeters, and especially of the so-called Transition-Edge Sensors (TES) on future satellites (for more information on future instrumentation see Paerels et al. 2008—Chap. 19, this issue). Compared to the current state-of-the-art CCD technology, this new type of detectors promises a factor of > 20 improvement in spectral resolution ($\lesssim$ 2–5 eV, depending on the optimisation of the detector). To reach this high spectral resolution, the TES detectors use the transition edge between normal conductivity and superconductivity that is present in a number of materials. At temperatures of ∼0.1 K, the conductivity of the absorbing material is very sensitive to the temperature. When an X-ray photon hits the detector, its temperature rises and the resistance of the material increases. By measuring the increase of the resistance, one can calculate with high accuracy the energy of the X-ray photon. The main effort today is to build an array of TES detectors, with imaging capabilities. Next major proposed missions with TES arrays on board are, amongst others, the *X-ray Evolving Universe Explorer* (*XEUS*), *Explorer of Diffuse emission and Gamma-ray burst Explosions* (*EDGE*), and *Constellation-X*.

The combination of the large effective area (∼5 m^2) and high spectral resolution ($\lesssim$2–5 eV) of *XEUS* will open up new possibilities in cluster abundance studies, with enormous improvement in the abundance determinations. In the left panel of Fig. 12 we show a simulated spectrum of the cluster Sérsic 159-03 (from de Plaa 2007), for the planned TES detector that will be part of the Narrow-Field Imager (NFI) on *XEUS*. For comparison also the model at the resolution of *XMM-Newton* EPIC is shown. With *XEUS* the abundances will be measured with an accuracy of $\sim 10^{-3}$ Solar, which is more than an order of magnitude improvement compared to EPIC. This accuracy might be sufficient to detect the contribution of the Population-III stars to the enrichment of the ICM. Also spectral lines from more elements will be resolved, which will enable better tests for supernova models.

The large collecting area of *XEUS* in combination with the CCD-type detectors of the Wide-Field Imager will be excellent to observe the spatial distribution of metals, which will help us to constrain the enrichment history of clusters. Simulated metallicity maps by Kapferer et al. (2006) for *Chandra*, *XMM-Newton*, and *XEUS* are shown in the right panel of Fig. 12.

The proposed *EDGE* satellite would also open up new possibilities in the chemical abundance studies. The combination of the planned long exposure times (of the order of few Ms), large field of view (1.4 degrees) and good angular resolution (Half Power Diameter of 15″) of its CCD type Wide-Field Imager, and the high spectral resolution (3 eV at 0.6 keV) of the

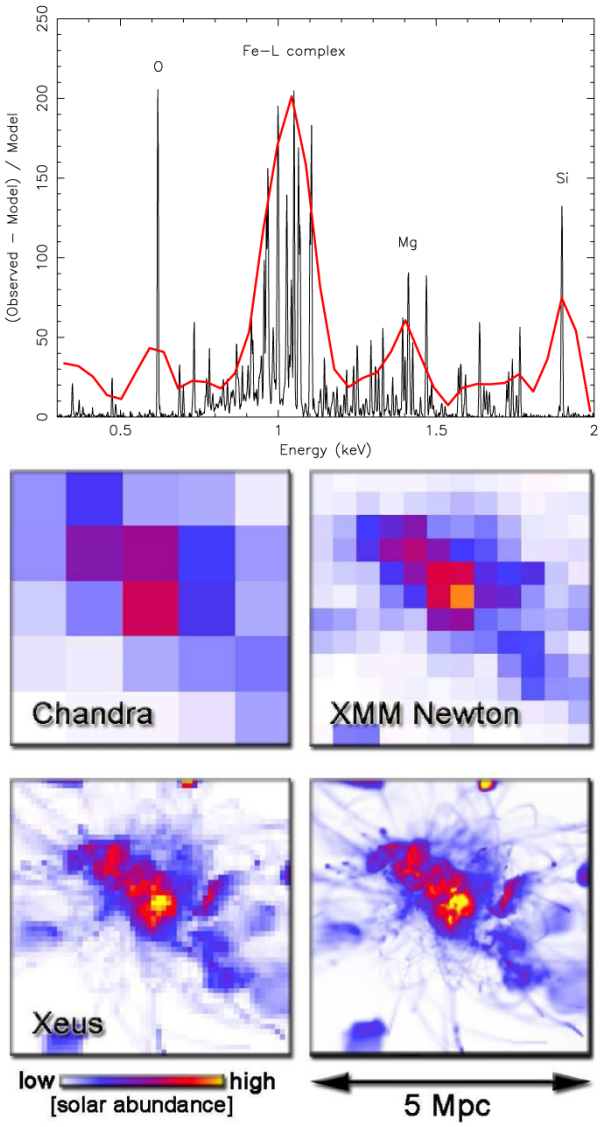


Fig. 12 *Top panel:* Simulated line spectrum of the inner 30″ of the cluster of galaxies Sérsic 159-03 for a 100 ks observation with the TES array planned for XEUS. The continuum model is subtracted. For comparison the EPIC spectrum of the same source is overplotted as the *thick line*. From de Plaa (2007). *Bottom panel*: Simulated X-ray metallicity maps as seen with *Chandra*, *XMM-Newton*, and *XEUS*, respectively. The *plot in the bottom right corner* shows the original simulation results. From Kapferer et al. (2006)

Wide-field Spectrometer promises new possibilities in the studies of metal abundances in clusters. *EDGE* would not only increase the number of detected spectral lines and improve the accuracy of the abundance measurements, but it would also enable us to accurately measure the chemical abundances and the spatial distribution of metals in the faint outskirts of clusters out to their virial radius, which is beyond the possibilities of current instruments. This would provide us with insights of crucial importance for the understanding of the chemical enrichment processes, nucleosynthesis, and of the metal budget of the Universe.

Acknowledgements The authors thank ISSI (Bern) for the support of the team "Non-virialised X-ray components in clusters of galaxies". We would like to thank Hans Böhringer and Alexis Finoguenov for reading the manuscript and providing comments. NW acknowledges support by the Marie Curie EARA Early Stage

Training visiting fellowship. The Netherlands Institute for Space Research is supported financially by NWO, the Netherlands Organization for Scientific Research. We thank Joop Schaye for his help with the core collapse supernova yields. RW acknowledges support from Marie Curie Excellence Grant MEXT-CT-2004-014112.

References

E. Anders, N. Grevesse, Geochim. Cosmochim. Acta **53**, 197 (1989)
M. Arnaud, R. Rothenflug, O. Boulade, L. Vigroux, E. Vangioni-Flam, Astron. Astrophys. **254**, 49 (1992)
C. Badenes, K.J. Borkowski, J.P. Hughes, U. Hwang, E. Bravo, Astrophys. J. **645**, 1373 (2006)
J. Bagchi, F. Durret, G.B. Lima Neto, S. Paul, Science **314**, 791 (2006)
I. Balestra, P. Tozzi, S. Ettori et al., Astron. Astrophys. **462**, 429 (2007)
W.H. Baumgartner, M. Loewenstein, D.J. Horner, R.F. Mushotzky, Astrophys. J. **620**, 680 (2005)
T. Bensby, S. Feltzing, Mon. Not. R. Astron. Soc. **367**, 1181 (2006)
H. Böhringer, E. Belsole, J. Kennea et al., Astron. Astrophys. **365**, L181 (2001)
H. Böhringer, K. Matsushita, E. Churazov, A. Finoguenov, Y. Ikebe, Astron. Astrophys. **416**, L21 (2004)
H. Böhringer, K. Matsushita, A. Finoguenov, Y. Xue, E. Churazov, Adv. Sp. Res. **36**, 677 (2005)
S. Borgani, D. Fabjan, L. Tornatore et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9322-7
D.A. Buote, Astrophys. J. **539**, 172 (2000a)
D.A. Buote, Mon. Not. R. Astron. Soc. **311**, 176 (2000b)
D.A. Buote, Astrophys. J. **574**, L135 (2002)
D.A. Buote, C.R. Canizares, Astrophys. J. **427**, 86 (1994)
D.A. Buote, A.C. Fabian, Mon. Not. R. Astron. Soc. **296**, 977 (1998)
D.A. Buote, A.D. Lewis, F. Brighenti, W.G. Mathews, Astrophys. J. **595**, 151 (2003)
D.A. Buote, F. Brighenti, W.G. Mathews, Astrophys. J. **607**, L91 (2004)
C.R. Canizares, G.W. Clark, T.H. Markert et al., Astrophys. J. **234**, L33 (1979)
C. Chiappini, F. Matteucci, G. Meynet, Astron. Astrophys. **410**, 257 (2003)
A. Chieffi, M. Limongi, Astrophys. J. **608**, 405 (2004)
G. Clementini, R.G. Gratton, E. Carretta, C. Sneden, Mon. Not. R. Astron. Soc. **302**, 22 (1999)
S. De Grandi, S. Molendi, Astrophys. J. **551**, 153 (2001)
S. De Grandi, S. Ettori, M. Longhetti, S. Molendi, Astron. Astrophys. **419**, 7 (2004)
J. de Plaa, Enrichment study of hot intra-cluster gas through X-ray spectroscopy. PhD thesis, Utrecht University, 2007
J. de Plaa, N. Werner, A.M. Bykov et al., Astron. Astrophys. **452**, 397 (2006)
J. de Plaa, N. Werner, J.A.M. Bleeker et al., Astron. Astrophys. **465**, 345 (2007)
M. Donahue, G.M. Voit, C.A. Scharf et al., Astrophys. J. **527**, 525 (1999)
R.A. Dupke, R.E. White, III, Astrophys. J. **528**, 139 (2000)
F. Durret, G.B. Lima Neto, W. Forman, Astron. Astrophys. **432**, 809 (2005)
S. Ettori, Mon. Not. R. Astron. Soc. **362**, 110 (2005)
S. Ettori, A.C. Fabian, Mon. Not. R. Astron. Soc. **305**, 834 (1999)
U. Feldman, Phys. Scr. **46**, 202 (1992)
A. Finoguenov, L.P. David, T.J. Ponman, Astrophys. J. **544**, 188 (2000)
A. Finoguenov, K. Matsushita, H. Böhringer, Y. Ikebe, M. Arnaud, Astron. Astrophys. **381**, 21 (2002)
A. Finoguenov, M.J. Henriksen, F. Miniati, U.G. Briel, C. Jones, Astrophys. J. **643**, 790 (2006)
Y. Fujita, N. Tawa, K. Hayashida et al., Publ. Astron. Soc. Jpn. (2007, in press). astro-ph/0705.2017
Y. Fukazawa, T. Ohashi, A.C. Fabian et al., Publ. Astron. Soc. Jpn. **46**, L55 (1994)
Y. Fukazawa, K. Makishima, T. Tamura et al., Publ. Astron. Soc. Jpn. **50**, 187 (1998)
F. Gastaldello, S. Molendi, Astrophys. J. **572**, 160 (2002)
R. Giacconi, G. Branduardi, U. Briel et al., Astrophys. J. **230**, 540 (1979)
N. Grevesse, A.J. Sauval, Space Sci. Rev. **85**, 161 (1998)
B. Gustafsson, T. Karlsson, E. Olsson, B. Edvardsson, N. Ryde, Astron. Astrophys. **342**, 426 (1999)
M. Hamuy, M.M. Phillips, N.B. Suntzeff et al., Astron. J. **112**, 2438 (1996)
P.J. Humphrey, D.A. Buote, Astrophys. J. **639**, 136 (2006)
P.J. Humphrey, D.A. Buote, C.R. Canizares, Astrophys. J. **617**, 1047 (2004)
V.D. Ivanov, M. Hamuy, P.A. Pinto, Astrophys. J. **542**, 588 (2000)
K. Iwamoto, F. Brachwitz, K. Nomoto et al., Astrophys. J. Suppl. Ser. **125**, 439 (1999)
J.S. Kaastra, An X-ray spectral code for optically thin plasmas. Internal SRON-Leiden Report, Updated Version 2.0, 1992
W. Kapferer, C. Ferrari, W. Domainko et al., Astron. Astrophys. **447**, 827 (2006)

W. Kapferer, T. Kronberger, J. Weratschnig, S. Schindler, Astron. Astrophys. **472**, 757 (2007a)
W. Kapferer, T. Kronberger, J. Weratschnig et al., Astron. Astrophys. **466**, 813 (2007b)
M. Kirsch, 2006. http://xmm.vilspa.esa.es/docs/documents/CAL-TN-0018.pdf
K. Koyama, H. Tsunemi, T. Dotani et al., Publ. Astron. Soc. Jpn. **59**, 23 (2007)
S.M. Lea, R. Mushotzky, S.S. Holt, Astrophys. J. **262**, 24 (1982)
D.A. Liedahl, A.L. Osterheld, W.H. Goldstein, Astrophys. J. **438**, L115 (1995)
Y.-T. Lin, J.J. Mohr, S.A. Stanford, Astrophys. J. **591**, 749 (2003)
K. Lodders, Astrophys. J. **591**, 1220 (2003)
A. Maeder, G. Meynet, Astron. Astrophys. **361**, 159 (2000)
F. Mannucci, M. Della Valle, N. Panagia, Mon. Not. R. Astron. Soc. **370**, 773 (2006)
K. Matsushita, A. Finoguenov, H. Böhringer, Astron. Astrophys. **401**, 443 (2003)
K. Matsushita, H. Böhringer, I. Takahashi, Y. Ikebe, Astron. Astrophys. **462**, 953 (2007a)
K. Matsushita, Y. Fukazawa, J.P. Hughes et al., Publ. Astron. Soc. Jpn. **59**, 327 (2007b)
F. Matteucci, M. Tosi, Mon. Not. R. Astron. Soc. **217**, 391 (1985)
B.J. Maughan, C. Jones, W. Forman, L. Van Speybroeck, Astrophys. J. Suppl. Ser. **174**, 117 (2008)
R. Mewe, E.H.B.M. Gronenschild, G.H.J. van den Oord, Astrophys. J. Suppl. Ser. **62**, 197 (1985)
R. Mewe, J.R. Lemen, G.H.J. van den Oord, Astrophys. J. Suppl. Ser. **65**, 511 (1986)
R. Mewe, J.S. Kaastra, D.A. Liedahl, Legacy **6**, 16 (1995)
G. Meynet, A. Maeder, Astron. Astrophys. **381**, L25 (2002)
R.J. Mitchell, J.L. Culhane, P.J.N. Davison, J.C. Ives, Mon. Not. R. Astron. Soc. **175**, 29P (1976)
K. Mitsuda, M. Bautz, H. Inoue et al., Publ. Astron. Soc. Jpn. **59**, 1 (2007)
S. Molendi, F. Gastaldello, Astron. Astrophys. **375**, L14 (2001)
R.F. Mushotzky, Physica Scripta **T7**, 157 (1984)
R.F. Mushotzky, M. Loewenstein, Astrophys. J. **481**, L63 (1997)
R.F. Mushotzky, P.J. Serlemitsos, E.A. Boldt, S.S. Holt, B.W. Smith, Astrophys. J. **225**, 21 (1978)
R.F. Mushotzky, S.S. Holt, E.A. Boldt, P.J. Serlemitsos, B.W. Smith, Astrophys. J. **244**, L47 (1981)
R. Mushotzky, M. Loewenstein, K.A. Arnaud et al., Astrophys. J. **466**, 686 (1996)
K. Nomoto, F.-K. Thielemann, K. Yokoi, Astrophys. J. **286**, 644 (1984)
K. Nomoto, N. Tominaga, H. Umeda, C. Kobayashi, K. Maeda, Nucl. Phys. A **777**, 424 (2006)
K. Nomoto, M. Hashimoto, T. Tsujimoto et al., Nucl. Phys. A **616**, 79 (1997)
P.E.J. Nulsen, G.C. Stewart, A.C. Fabian et al., Mon. Not. R. Astron. Soc. **199**, 1089 (1982)
E. O'Sullivan, J.M. Vrtilek, J.C. Kempner, L.P. David, J.C. Houck, Mon. Not. R. Astron. Soc. **357**, 1134 (2005)
F.N. Owen, J.A. Eilek, N.E. Kassim, Astrophys. J. **543**, 611 (2000)
F. Paerels, J.S. Kaastra, T. Ohashi et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9323-6
J.R. Peterson, S.M. Kahn, F.B.S. Paerels et al., Astrophys. J. **590**, 207 (2003)
G.W. Pratt, H. Böhringer, J.H. Croston et al., Astron. Astrophys. **461**, 71 (2007)
P. Rebusco, E. Churazov, H. Böhringer, W. Forman, Mon. Not. R. Astron. Soc. **359**, 1041 (2005)
P. Rebusco, E. Churazov, H. Böhringer, W. Forman, Mon. Not. R. Astron. Soc. **372**, 1840 (2006)
A. Renzini, Astrophys. J, **488**, 35 (1997)
A. Renzini, in *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, ed. by J.S. Mulchaey, A. Dressler, A. Oemler, 2004, p. 260
E. Rizza, J.O. Burns, M.J. Ledlow et al., Mon. Not. R. Astron. Soc. **301**, 328 (1998)
J.S. Sanders, A.C. Fabian, Mon. Not. R. Astron. Soc. **331**, 273 (2002)
J.S. Sanders, A.C. Fabian, Mon. Not. R. Astron. Soc. **371**, 1483 (2006)
J.S. Sanders, A.C. Fabian, Mon. Not. R. Astron. Soc. **381**, 1381 (2007)
J.S. Sanders, A.C. Fabian, S.W. Allen, R.W. Schmidt, Mon. Not. R. Astron. Soc. **349**, 952 (2004)
C.L. Sarazin, *X-Ray Emissions from Clusters of Galaxies* (Cambridge University Press, Cambridge, 1988)
K. Sato, K. Matsushita, Y. Ishisaki et al., Publ. Astron. Soc. Jpn. (2007a, in press). astro-ph/0707.4342
K. Sato, K. Tokoi, K. Matsushita et al., Astrophys. J. **667**, 41 (2007b)
K. Sato, N.Y. Yamasaki, M. Ishida et al., Publ. Astron. Soc. Jpn. **59**, 299 (2007c)
J.L. Sauvageot, E. Belsole, G.W. Pratt, Astron. Astrophys. **444**, 673 (2005)
S. Schindler, A. Diaferio, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9321-8
P.J. Serlemitsos, B.W. Smith, E.A. Boldt, S.S. Holt, J.H. Swank, Astrophys. J. **211**, L63 (1977)
P.J. Serlemitsos, Y. Soong, K.-W. Chan et al., Publ. Astron. Soc. Jpn. **59**, 9 (2007)
J.R. Shi, G. Zhao, Y.Q. Chen, Astron. Astrophys. **381**, 982 (2002)
A. Simionescu, N. Werner, A. Finoguenov, H. Böhringer, M. Brüggen, Astron. Astrophys. (2007, in press). astro-ph/0709.4499
R.K. Smith, N.S. Brickhouse, D.A. Liedahl, J.C. Raymond, Astrophys. J. **556**, L91 (2001)
T. Tamura, J.A.M. Bleeker, J.S. Kaastra, C. Ferrigno, S. Molendi, Astron. Astrophys. **379**, 107 (2001)

T. Tamura, J.S. Kaastra, J.W.A. den Herder, J.A.M. Bleeker, J.R. Peterson, Astron. Astrophys. **420**, 135 (2004)
L.K. Townsley, P.S. Broos, J.A. Nousek, G.P. Garmire, Nucl. Instr. Meth. Phys. Res. A **486**, 751 (2002)
A. Vikhlinin, M. Markevitch, S.S. Murray et al., Astrophys. J. **628**, 655 (2005)
N. Werner, H. Böhringer, J.S. Kaastra et al., Astron. Astrophys. **459**, 353 (2006a)
N. Werner, J. de Plaa, J.S. Kaastra et al., Astron. Astrophys. **449**, 475 (2006b)
R.E. White III, C.S.R. Day, I. Hatsukade, J.P. Hughes, Astrophys. J. **433**, 583 (1994)
S.E. Woosley, T.A. Weaver, Astrophys. J. Suppl. Ser. **101**, 181 (1995)
H. Xu, S.M. Kahn, J.R. Peterson et al., Astrophys. J. **579**, 600 (2002)

Chapter 17
Metal Enrichment Processes

S. Schindler · A. Diaferio

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9321-8

Abstract There are many processes that can transport gas from the galaxies to their environment and enrich the environment in this way with metals. These metal enrichment processes have a large influence on the evolution of both the galaxies and their environment. Various processes can contribute to the gas transfer: ram-pressure stripping, galactic winds, AGN outflows, galaxy-galaxy interactions and others. We review their observational evidence, corresponding simulations, their efficiencies, and their time scales as far as they are known to date. It seems that all processes can contribute to the enrichment. There is not a single process that always dominates the enrichment, because the efficiencies of the processes vary strongly with galaxy and environmental properties.

Keywords Galaxies: clusters: general · ISM: jets and outflows · Galaxies: ISM · Galaxies: interaction

1 Introduction

The gas between the galaxies in a cluster—the Intra-Cluster Medium (ICM)—does not only contain primordial elements, but also a considerable amount of heavy elements like Fe, Si, S, or O (see Werner et al. 2008—Chap. 16, this issue) resulting in metallicities around 0.5 in Solar units and sometimes even higher values. A large fraction ($\approx$ 15–20%) of the total mass of a cluster is in the ICM, whereas the galaxies contribute a substantially smaller

S. Schindler (✉)
Institute for Astro- and Particle Physics, University of Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria
e-mail: Sabine.Schindler@uibk.ac.at

A. Diaferio
Dipartimento di Fisica Generale "Amedeo Avogadro", Università degli Studi di Torino, Via P. Giuria 1, 10125, Torino, Italy

A. Diaferio
Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy

fraction (3–5%), and the rest is dark matter. It follows that there is more mass in metals in the ICM than in all the galaxies of a cluster. This means that a lot of metals must have been transported from the galaxies into the ICM. This gas transfer affects the evolution of galaxies and of galaxy clusters. When galaxies lose their gas, the star formation rate decreases and consequently the properties of the galaxies change. Depending on the time and the efficiency of the gas removal the evolution of the galaxies is more or less affected. Therefore it is important to know when, where and how the gas transport takes place.

Various processes are discussed that can contribute to the metal enrichment—some depend only on internal properties of the galaxies, others on the environment or the combination of both. We review here several enrichment processes: ram-pressure stripping, galactic winds, AGN outflows, galaxy-galaxy interactions and the effect of an intra-cluster stellar population. Please note that this list is certainly not complete and further processes might also contribute a small fraction to the metal enrichment of the ICM. Furthermore, some processes influence each other, which makes the picture even more complicated.

For several of the processes not only observational evidence exists, but also numerical simulations have been performed. We review here both aspects.

2 Ram-Pressure Stripping

A galaxy passing through the ICM feels an external pressure. This pressure depends on the ICM density $\rho_{\rm ICM}$ and the relative velocity $v_{\rm rel}$ of the galaxy and the ICM. Gunn and Gott (1972) suggested this process already many years ago. They also gave a frequently used prescription for the radius r beyond which the gas of a galaxy is stripped, depending on the ram pressure and the galactic gravitational restoring force. The implicit condition on r reads

$$p_{\rm ram} = \rho_{\rm ICM} v_{\rm rel}^2 > 2\pi G \sigma_{\rm star}(r) \sigma_{\rm gas}(r) \tag{1}$$

with $p_{\rm ram}$ being the ram pressure, G the gravitational constant, $\sigma_{\rm star}$ the stellar surface density, $\sigma_{\rm gas}$ the surface mass density of the galactic gas.

2.1 Observations

Nowadays the process of ram-pressure stripping receives more and more attention. There is now much observational evidence of stripped galaxies. In the Virgo cluster several examples of spiral galaxies affected by ram-pressure stripping have been found by H I observations (Cayatte et al. 1990; Veilleux et al. 1999; Vollmer et al. 1999, 2004a, 2004b; Vollmer 2003; Kenney et al. 2004; Koopmann and Kenney 2004; Crowl et al. 2005, see Fig. 1). Furthermore in Virgo elliptical galaxies stripping features have been discovered (e.g. Rangarajan et al. 1995; Lucero et al. 2005; Machacek et al. 2006a). Also in Coma and other clusters and groups evidence for ram-pressure stripping has been found (Bravo-Alfaro et al. 2000, 2001; Kemp et al. 2005; Rasmussen et al. 2006; Levy et al. 2007). Deficiency of H I has been reported as evidence for ram-pressure stripping (Vollmer and Huchtmeier 2007). It is even possible that the H I plume of the galaxy NGC 4388, that extends to more than 100 kpc, is a ram-pressure stripping feature (Oosterloo and van Gorkom 2005). In the galaxy NGC 7619 stripping features have been found showing a high metallicity in the gas tail behind the galaxy (Kim et al. 2007). The galaxy ESO 137-001 in the cluster Abell 3627 shows a long tail with several star formation regions suggesting that ram-pressure triggers star formation not only within the galaxy but also in the stripped material (Sun et al. 2007). Recently ram-pressure features have also been found in distant clusters (Cortese et al. 2007).

For a review on ram-pressure stripping and H I deficiency see van Gorkom (2003).

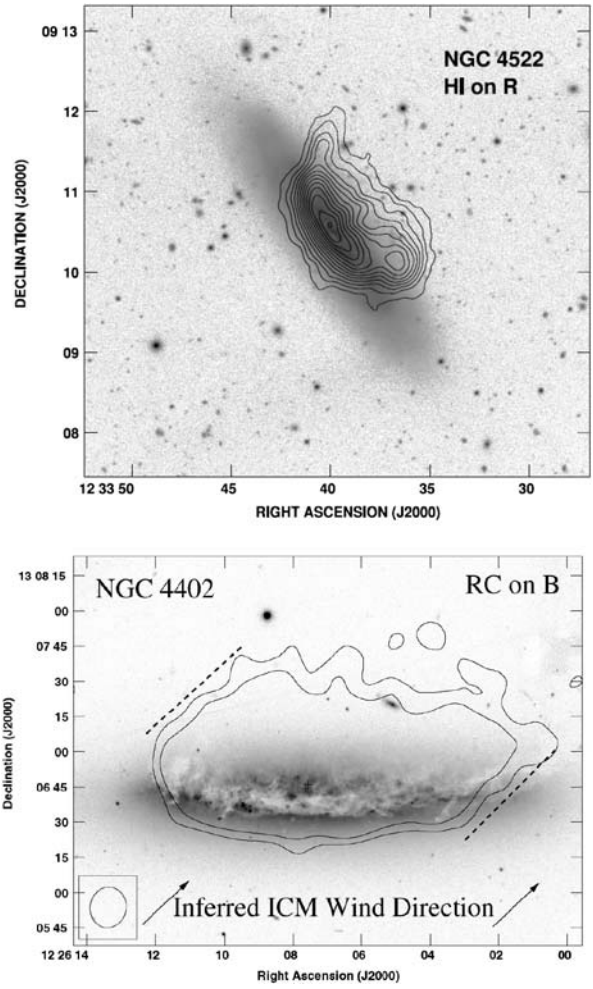


Fig. 1 Two examples of galaxies affected by ram-pressure stripping in the Virgo cluster. *Top*: H I gas (contours) and stellar light (grey scale, R-band) of NGC 4522. While the stellar distribution looks undisturbed the gas is bent back due to ram-pressure stripping (1.4 GHz radio continuum contours on B-band, from Kenney et al. 2004). *Bottom*: Similar features in NGC 4402 (from Crowl et al. 2005)

2.2 Simulations

As ram-pressure stripping is such a common process, there are many simulations in which the stripping process of galaxies was calculated for different types of galaxies: spirals, ellipticals and dwarfs (Abadi et al. 1999; Quilis et al. 2000; Mori and Burkert 2000; Toniazzo and Schindler 2001; Schulz and Struck 2001; Vollmer et al. 2001, 2006; Hidaka and Sofue 2002; Bekki and Couch 2003; Otmianowska-Mazur and Vollmer 2003; Acreman et al. 2003; Marcolini et al. 2003; Roediger and Hensler 2005; Roediger and Brüggen 2006; Roediger et al. 2006; Mayer et al. 2006, see Fig. 2). The simulations confirm that the process is acting in the expected way. Starting from the outer parts of the galaxy gas is stripped off. Part of this gas is not bound to the galaxies anymore and left in a wide (fragmenting) tail behind the galaxy. Recently even the increase of star formation in and behind the galaxy caused by ram-pressure stripping has been found in simulations (Kronberger et al. 2008a; Kapferer et al. 2008). Brüggen and De Lucia (2008) found that more than half of the cluster galaxies have experienced ram-pressure stripping and hence a considerable fraction of galaxies in a cluster at the present epoch has suffered a substantial gas loss.

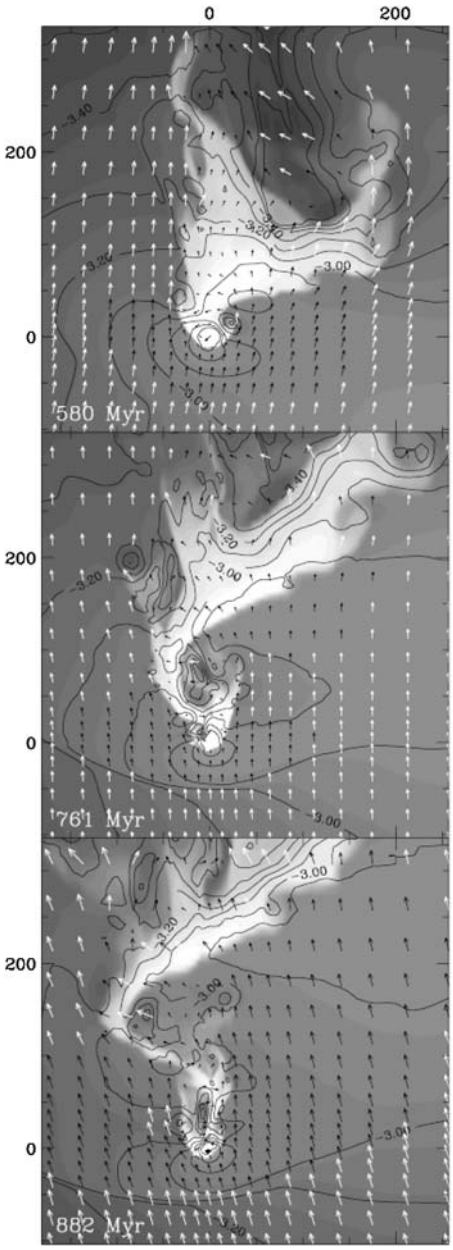


Fig. 2 Simulation of a galaxy in the process of ram-pressure stripping with gas density (*grey scale*, ranging from 4×10^{-29} to 1×10^{-27} g cm^{-3}), pressure (*black contours*, logarithmically spaced from $\log(p) = -3.6$ to $\log(p) = -2.6$ in units of 10^{-24} keV cm^{-3}) and velocity vectors (white when Mach number > 1, and black otherwise) at three different times. A cut through a 3D simulations is shown with coordinates in kpc. Due to the ram pressure of the ICM the galaxy loses more and more of its gas (from Toniazzo and Schindler 2001)

An analytical model has been developed to describe ram-pressure stripping for galaxies of different morphologies in different environments (Hester 2006). It describes the stripping of a satellite galaxy's outer H I disk and hot galactic halo.

It was tested with simulations whether the simple, widely used criterion by Gunn and Gott (1972) is a good estimate for the mass loss. Generally it is found that the criterion is a good estimate for the mass loss (Roediger and Brüggen 2007; Kronberger et al. 2008a) when simulations and analytic estimates are compared for the same conditions (see Jachym

et al. 2007 for a comparison with different conditions)—a quite surprising result given the simple assumptions of the criterion, that do not even take into account dark matter.

3 Galactic Winds

Already many years ago galactic winds were suggested as a possible gas transfer mechanism (De Young 1978). Many supernova explosions provide large amounts of thermal energy, which can drive an outflow from a galaxy (see reviews by Heckman 2003 and Veilleux et al. 2005). A correlation between starburst galaxies and wind is well established through the finding of hot gas around starburst galaxies (e.g. Dahlem et al. 1998).

Spectacular examples of such winds are seen in the galaxies M 82 (Lynds and Sandage 1963) and NGC 253 (Demoulin and Burbidge 1970).

The outflows consist of a complex multi-phase medium of cool, warm and hot gas (see e.g. the *Chandra* observation of NGC 4631, Wang et al. 2001, Fig. 3). The morphologies of the optical emission-line gas and the X-ray emission as observed with *Chandra* have been found to be quite similar (Strickland et al. 2002; Cecil et al. 2002). Such correlations can be used to understand the interaction between the gas in the bubbles and the interstellar medium (ISM). It was found that the accelerated ISM can reach high velocities of several hundred $\mathrm{km\,s^{-1}}$ (Heckman et al. 2000; Rupke et al. 2002).

With these winds also metals are transported into the ICM. The amount of metals depends on various galaxy parameters, like the total mass of the galaxy or the disc scale length, and on the environmental conditions: e.g. in the centre of massive clusters the pressure of the ICM can suppress the winds (Kapferer et al. 2006). This suppression typically takes place for ICM pressures above $(0.7\text{–}1) \times 10^{12}\,\mathrm{dyne\,cm^{-2}}$.

Martin (1999) gives an often used recipe for simulations: the mass outflow rate $\dot{M}$ is proportional to the star formation rate SFR:

$$\dot{M} = \epsilon \mathrm{SFR} \tag{2}$$

with ϵ being typically in the range of 1–3. By comparing different techniques Heckman (2003) also finds that the outflow rate is of the order of the star formation rate. The SFR can

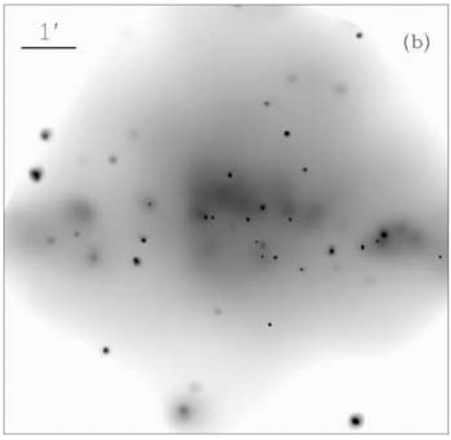


Fig. 3 *Chandra* observation of the edge-on spiral galaxy NGC 4631. It shows the presence of a giant diffuse X-ray emitting corona. The corona has a temperature of $(2\text{–}7) \times 10^6$ K and extends as far as 8 kpc away from the galactic plane (from Wang et al. 2001)

be estimated from observations, e.g. from far-infrared luminosities $L_{\rm FIR}$

$$\frac{\rm SFR}{1\ {\rm M}_\odot\ {\rm yr}^{-1}} = \frac{L_{\rm FIR}}{5.8\times 10^9\ {\rm L}_\odot} \tag{3}$$

(Kennicutt 1998). Another way to estimate the SFR is to use the tight relation between the SFR and the surface density of the gas $\sigma_{\rm gas}$

$$\Sigma_{\rm SFR} \propto \sigma_{\rm gas}^N \tag{4}$$

(Schmidt 1959) with $\Sigma_{\rm SFR}$ being the surface density of the SFR and the index N having measured values between 1 and 2. Only at densities below a critical threshold value the SFR is almost completely suppressed (Kennicutt 1989). Alternatively the dynamical time t_* can be included

$$\Sigma_{\rm SFR} \propto \frac{\Sigma_{\rm gas}}{t_*} \tag{5}$$

with the dynamical time t_* being the local orbital timescale of the disk (Kennicutt 1998). For hydrodynamic simulations this has been extended to include the fraction of stars lost by supernova explosions by Springel and Hernquist (2003)

$$\frac{{\rm d}\rho_*}{{\rm d}t} = (1-\beta)\frac{{\rm d}\rho_{\rm c}}{{\rm d}t} \tag{6}$$

with ρ_* being the density of stars, $\rho_{\rm c}$ being the cold gas density in the disk and β being the fraction of stars lost by supernova explosions. For a typical initial mass function and a mass threshold of 8 $\rm M_\odot$ for the supernovae a $\beta = 0.1$ is used. These are of course only statistical estimates.

Other attempts to quantify the outflow rate take into account physical parameters like those describing the galaxy's gravitational potential and the effect of cosmic rays (Breitschwerdt et al. 1991). Using the Bernoulli equation, Kronberger et al. (2008b) recently derived an analytic approximation for the mass loss due to thermally driven galactic winds. The mass loss per unit area at a given position of the galactic disc reads

$$\dot{M} = \rho_0 u_0 = \rho_0 \sqrt{v_{\rm esc}^2 + 2\Phi_0 - \frac{5}{\gamma}c_0^2}, \tag{7}$$

with ρ_0 being the gas mass density, u_0 the bulk velocity of the gas, Φ_0 the gravitational potential, c_0 the sound speed (all four quantities at the given position), $v_{\rm esc}$ the escape velocity, and γ the adiabatic index of the thermal plasma. Hydrodynamic simulations of outflows have also been performed (Tenorio-Tagle and Munoz-Tunon 1998; Strickland and Stevens 2000).

Starbursts with subsequent winds can also be caused by cluster mergers (Ferrari et al. 2003, 2005, 2006), because in such mergers the gas is compressed and shock waves and cold fronts, which trigger star formation, are produced (Evrard 1991; Caldwell et al. 1993; Wang et al. 1997; Owen et al. 1999; Moss and Whittle 2000; Bekki and Couch 2003).

4 Galaxy-Galaxy Interaction

Another possible mechanism for removing material—gas and stars—from galaxies is the interaction between the galaxies (e.g. Clemens et al. 2000; Mihos et al. 2005, see Figs. 4,

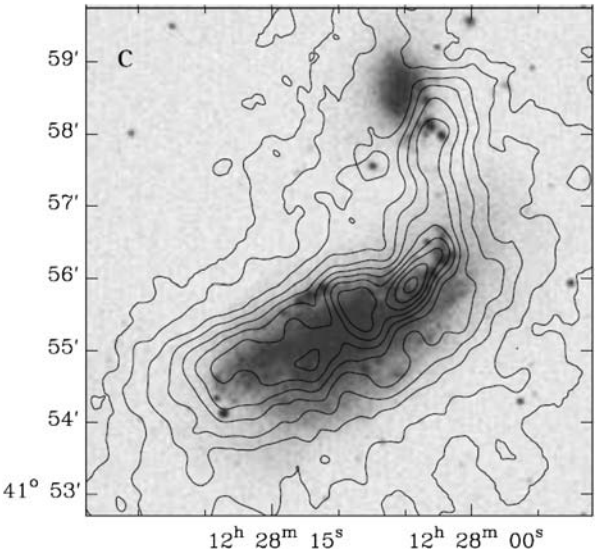


Fig. 4 Image of the interacting galaxies NGC 4490 / NGC 4485 in H I (*contours*) and optical R band (*grey scale*). Some of the gas is lost due to the interaction of the galaxies (from Clemens et al. 2000)

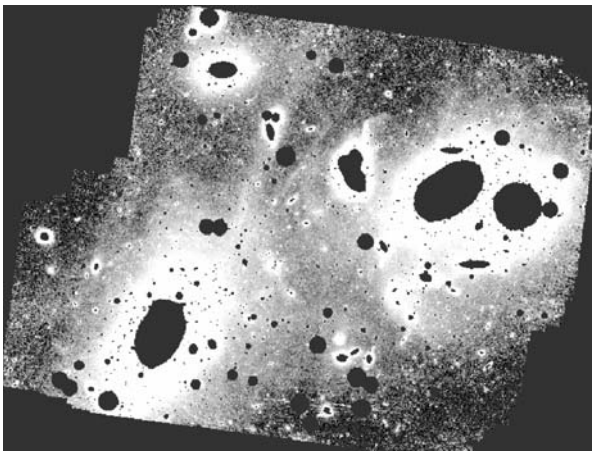


Fig. 5 Very deep observation of the core of the Virgo cluster. Diffuse light is visible between the galaxies which results from stars that have been expelled from the galaxies due to interactions between them (from Mihos et al. 2005)

5 and 6). While the direct stripping effect is mostly not very efficient in clusters due to the short interaction times, the close passage of another galaxy (sometimes called galaxy harassment) can trigger a star burst (Barnes and Hernquist 1992; Moore et al. 1996; Bekki 1999), which subsequently can lead to a galactic wind (Kapferer et al. 2005). But there can be a competing effect: the ISM might be stripped off immediately by ram pressure (Fujita et al. 1999; Heinz et al. 2003) and hence the star formation rate could drop. In any case ISM would be removed from the galaxies.

Simulations of interactions between galaxies containing an AGN show a complex interplay between star formation and the activity of the AGN itself (Springel et al. 2005).

In order to estimate how likely such interaction events are the number of encounters and mergers needs to be assessed. The number of close encounters that a galaxy experiences within $t_{\mathrm{H}} = 10^{10}$ years was estimated by Gnedin (2003) in the following way. He assumed a galaxy of size $R_{\mathrm{g}} = 10$ kpc, a virialised cluster with a one-dimensional velocity dispersion

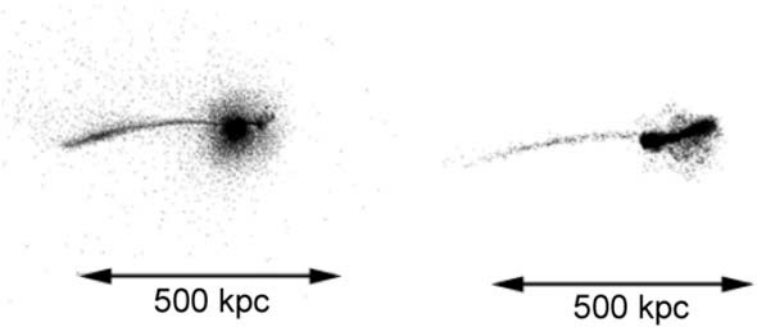


Fig. 6 Simulation of the interaction between two galaxies: distribution of gas (*left*) and stars (*right*). Due to the interaction gas and stars can be expelled out to distances of hundreds of kpc (from Kapferer et al. 2005)

of $\sigma_{\rm cl} = 1000\,{\rm km\,s^{-1}}$, a virial radius of $R_{\rm cl} = 1$ Mpc and $N_{\rm g} = 1000$ galaxies uniformly distributed within this radius. With a relative velocity of $\sqrt{2}\sigma_{\rm cl}$ and neglecting the gravitational focusing, he finds

$$N_{\rm enc} \approx \frac{N_{\rm g}}{(4\pi/3)R_{\rm cl}^3}\pi R_{\rm g}^2\sqrt{2}\sigma_{\rm cl}t_{\rm H} \approx 1, \tag{8}$$

i.e. a galaxy is expected to encounter one other galaxy over the course of its evolution. Even though the assumption is very simplifying one sees that an encounter is a relatively frequent event in a cluster. In contrast to this, Gnedin (2003) estimated for the probability to merge with another galaxy

$$P_{\rm mer} \approx N_{\rm enc}\left(\frac{\sigma_{\rm g}}{\sigma_{\rm cl}}\right)^4 \approx 10^{-3} \tag{9}$$

with a galactic velocity dispersion $\sigma_{\rm g} = 200\,{\rm km\,s^{-1}}$. Hence an actual merger is an unlikely event.

Tidal interactions and merging between galaxies are highly non-linear phenomena that can be partly handled analytically (e.g., Binney and Tremaine 1987, Chap. 7). Numerical simulations, however, are required for accurate estimates of mass loss rates and the morphological modification of galaxies.

5 AGN Outflows

We discuss two types of outflows from AGN: jets and winds-like outflows. There is much observational evidence for AGN jets interacting with the ICM—not only radio jets but also cavities in the ICM found in X-rays (e.g. Blanton et al. 2001; McNamara et al. 2001; Schindler et al. 2001; Heinz et al. 2002; Choi et al. 2004; Fabian et al. 2006; McNamara and Nulsen 2007), in which the pressure of the relativistic particles of the jet has pushed away the ICM. The jets consisting of relativistic particles can entrain some of the surrounding metal-rich gas (De Young 1986).

As the jet-ICM interaction can have an effect on both the energetics and the metal enrichment of the ICM, several groups have calculated this process. Many simulations for the energy transfer have been performed (Zhang et al. 1999; Churazov et al. 2001; Brüggen and Kaiser 2002; Nulsen et al. 2002; Krause and Camenzind 2003; Heinz 2003;

Beall et al. 2004, 2006; Dalla Vecchia et al. 2004; Zanni et al. 2005; Sijacki and Springel 2006; Heinz et al. 2006) while only few have attempted to calculate the metal enrichment due to the entrainment by jets (Heath et al. 2007; Moll et al. 2007). These simulations found that jets can both heat and enrich the ICM considerably. Another type of simulations calculated the metallicity distribution due to bubble-induced motions coming from a single AGN in the cluster centre (Roediger et al. 2007). It was found that in this case the metallicity distribution is very elongated along the direction of the motion of the bubbles.

Also for wind-like outflows there is some observational evidence. Blue-shifted absorption lines have been observed in UV and X-rays (Crenshaw et al. 2003). Also from X-ray imaging evidence for nuclear outflows has been found (Machacek et al. 2006b). There are hints for a high metallicity of a few times solar (Hamann et al. 2001; Hasinger et al. 2002), for high velocities of several thousands or several ten thousands of $\mathrm{km\,s^{-1}}$ (Chartas et al. 2002, 2003; Pounds et al. 2003a, 2003b; Reeves et al. 2003; O'Brian et al. 2005; Dasgupta et al. 2005; Gabel et al. 2006) and for considerable mass outflow rates (Crenshaw et al. 2003; Veilleux et al. 2005). The outflows can be quite strong, e.g. several 10^9 $\mathrm{M_\odot}$ with kinetic energies around 10^{60} erg expelled over the AGN live time of 10^7 years as estimated from spectroscopic studies (Nesvadba et al. 2006).

In some galaxies the winds are not only driven by repeated supernova explosions but also the AGN are contributing to the energy necessary for the wind (see Sect. 3).

6 Intra-Cluster Stellar Population

There is increasing evidence for a population of stars in the space between the galaxies in a cluster (Bernstein et al. 1995; Gonzalez et al. 2000, 2005; Gerhard et al. 2002, 2005; Gal-Yam et al. 2003; Arnaboldi et al. 2004; Cortese et al. 2004; Feldmeier et al. 2004; Ryan-Weber et al. 2004; Adami et al. 2005; Zibetti et al. 2005; Krick and Bernstein 2007). Depending on the mass of the cluster the fraction of intra-cluster stars (= ratio of number of stars between galaxies to total number of stars) can be as high as 10–50% with the higher fraction being in more massive clusters. This stellar population can originate from stripping of stars from galaxies due to tidal interaction (Cypriano et al. 2006), can be expelled during mergers and the formation of massive galaxies (Murante et al. 2007; Kapferer et al. 2005, Fig. 6) or can have multiple origins (Williams et al. 2007). Simulations show that intra-cluster stars should be ubiquitous in galaxy clusters (Willman et al. 2004) and their numbers should generally increase with time (Rudick et al. 2006). A link between the growth of the brightest cluster galaxy and the intra-cluster light was reported by Zibetti et al. (2005).

When these stars explode as supernovae (mainly type Ia, as it takes a while for the stars to travel away from the galaxies) they can enrich the ICM very efficiently because there is no ISM pressure around them to confine the metals (Domainko et al. 2004; Zaritzsky et al. 2004; Lin and Mohr 2004; Dado et al. 2007).

Considerably more frequent than supernova Ia explosions are their progenitors—the recurrent novae. With about 10^{-4} $\mathrm{M_\odot}$ outflow per nova event and typically super-solar abundances (up to ten times Solar, Gehrz et al. 1998), novae could also contribute to the metal enrichment of the ICM if they have been expelled previously from the galaxies.

A fraction of the AGB stars are also expected to be between the galaxies. These stars have a considerable mass loss with metallicities of about Solar abundances with slightly enhanced abundances of CNO elements (Wheeler et al. 1989; Zijlstra 2006; van den Hoek and Groenewegen 1997; Busso et al. 2001; Nordström 2003). As the ratio of planetary nebulae (PNe) to AGB stars is well studied in statistical studies of PNe in the ICM (Feldmeier et al. 1998;

Theuns and Warren 1997; Arnaboldi et al. 2003) this ratio may be used to estimate the number of AGB stars.

In conclusion the population of stars should also be considered for the enrichment processes in the ICM—even far away from galaxies.

7 Which of These Processes Are Important for the ICM Enrichment?

Several groups have addressed this question already many years ago. David et al. (1991) proposed the first models taking into account the effects of galactic winds on the ICM enrichment. They found that the results depend sensitively on their input parameters: the initial mass function, the adopted supernova rate and the primordial mass fraction of the ICM.

The first 3D simulations calculating the full gas dynamics and the effects of winds on cluster scales were performed by Metzler and Evrard (1994, 1997). They concluded that winds can account for the observed metal abundances in the ICM, but they found strong metallicity gradients (almost a factor of ten between cluster centre and virial radius) which are not in agreement with observations. Gnedin (1998) took into account not only galactic winds, but also galaxy-galaxy interactions and concluded that most of the metals are ejected by galaxy mergers. In contrast to this result Aguirre et al. (2001) found that galaxy-galaxy interactions and ram-pressure stripping are of minor importance while galactic winds dominate the metal enrichment of the ICM.

That these early results disagree so much is probably due to the large range of scales that is involved. On the one hand the whole cluster with its infall region has to be simulated, on the other hand processes within galaxies or even within the active core of a galaxy are important. It is not possible to calculate all of this accurately in one type of simulation and therefore new methods have been developed.

Recently several simulations for the enrichment have been performed. They calculate the exact composition and evolution of the ISM by varying the initial mass function and the yields of supernova explosions (see Borgani et al. 2008—Chap. 18, this issue), but it is not distinguished by which process the enriched gas is transported into the ICM. Specially for the transport processes a new simulation method has been developed, in which N-body/hydrodynamic simulations with mesh refinement including a semi-analytical method have been combined with separate descriptions of the various enrichment processes, which can be switched on and off individually (Schindler et al. 2005).

The results obtained with this method show an inhomogeneous distribution of the metals independent of the enrichment processes (Schindler et al. 2005; Domainko et al. 2006; Kapferer et al. 2006; Moll et al. 2007, see Fig. 7). These results are in very good agreement with the observed metallicity maps (see Werner et al. 2008—Chap. 16, this issue). The gas lost by the galaxies is obviously not mixed immediately with the ICM. There are usually several maxima visible in the metallicity distribution, which are not necessarily associated with the cluster centre. The maxima are typically at places where galaxies just have lost a lot of gas to ICM of low density, mostly due to star bursts. The metallicities vary locally between 0 and 4 times Solar.

A detailed comparison between the two enrichment mechanisms—winds and ram-pressure stripping—revealed that these two processes yield different metal distributions (see Fig. 7) and a different time dependence of the enrichment (Kapferer et al. 2007a; Rasia et al. 2007). The ram-pressure stripped gas is more centrally concentrated. The reason for this is that the ICM density as well as the galaxies velocities are larger in the cluster centre, so that ram-pressure stripping is very efficient there. Galactic winds, however, can

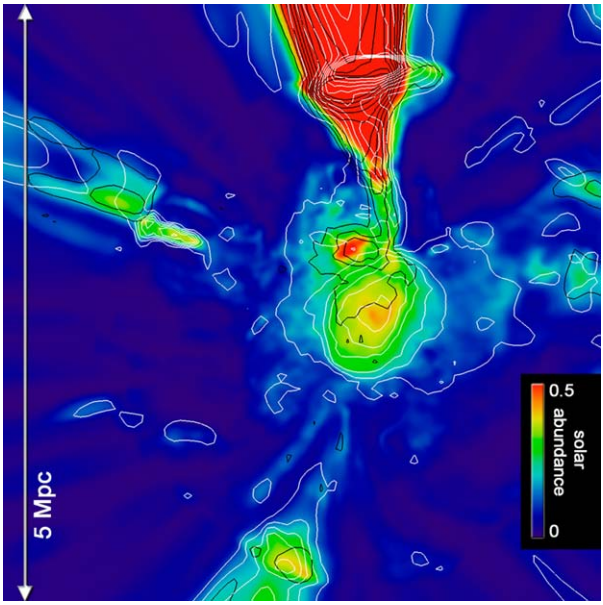


Fig. 7 Simulated metallicity map, i.e. an X-ray emission weighted, projected metal distribution. The high metallicity region at the top is caused by a group of galaxies with recent starburst. Overlaid contours indicate the origin of the metals: ram-pressure stripping (*white*) and galactic winds (*black*) (adopted from Kapferer et al. 2007a)

be suppressed by the high pressure of the ICM in the centre (Kapferer et al. 2006), so that in massive clusters galactic winds do hardly contribute to the central enrichment. The resulting radial metal profiles are correspondingly relatively flat for galactic winds and steep for ram-pressure stripping. When both processes are taken into account they are in good agreement with the observed profiles (see also Borgani et al. 2008—Chap. 18, this issue).

The time scales for the enrichment are also different for the two processes (Kapferer et al. 2007a). The mass loss of galaxies due to winds is larger at high redshifts. Between redshifts 2 and 1 ram-pressure stripping becomes more important for the mass loss and it is by far more efficient at low redshift (see Fig. 8). The reason is that on the one hand galactic winds become weaker because the star formation rate decreases and on the other hand ram-pressure stripping becomes stronger because clusters with ICM have formed/are forming, which is interacting with the galaxies. In total the mass loss due to ram-pressure stripping is usually larger than the mass loss due to winds, in some cases up to a factor of three.

Generally it is very hard to provide numbers for the relative efficiencies of the various processes as the efficiencies depend strongly on the properties of the clusters. In a massive or in a merger cluster, for example, ram-pressure stripping is very efficient.

The simulated metallicities can be converted to artificial X-ray metallicities, metallicity profiles, metallicity maps and metallicity evolution. There is in general a good agreement between these quantities derived from simulation and observation (Kapferer et al. 2007b). The metallicity values are in the right range and the spatial distribution and the evolution are in good agreement with the observations. Also the evolution of the metallicity since $z = 1$ found in observations (Balestra et al. 2007; Maughan et al. 2008) can be reproduced by the simulations. Of course there is a large scatter in all these quantities, because they vary very much from cluster to cluster both in simulations and observations.

Summarising, from the comparison of observations with simulations it seems clear that several processes are involved in the metal enrichment and none of them can be ruled out immediately as being not efficient enough. The processes can also influence each other (e.g. AGN outflows can enhance an existing galactic wind or one process can suppress another

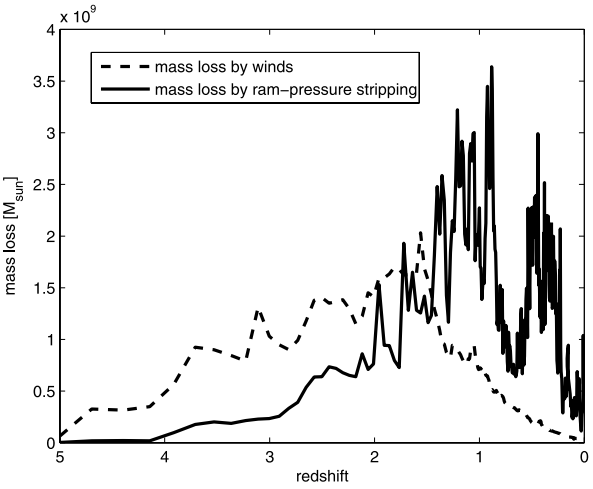


Fig. 8 Mass loss of the galaxies in a simulation taking into account mass loss due to galactic winds (*dashed line*) and mass loss due to ram-pressure stripping (*solid line*) at different redshifts (adopted from Kapferer et al. 2007a)

one). Obviously the interaction between galaxies and the ICM is a very complex issue. In order to know what is really going on at the transition between galaxies and ICM many more observations and simulations are needed.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialised X-ray components in clusters of galaxies". We thank Wolfgang Kapferer and Thomas Kronberger for useful discussions. The authors acknowledge financial support by the Austrian Science Foundation (FWF) through grants P18523-N16 and P19300-N16, by the Tiroler Wissenschaftsfonds and through the UniInfrastrukturprogramm 2005/06 by the BMWF. Partial support from the PRIN2006 grant "Costituenti fondamentali dell'Universo" of the Italian Ministry of University and Scientific Research and from the INFN grant PD51 is also gratefully acknowledged.

References

M.G. Abadi, B. Moore, R.G. Bower, Mon. Not. R. Astron. Soc. **308**, 947 (1999)
D.M. Acreman, I.R. Stevens, T.J. Ponman, I. Sakelliou, Mon. Not. R. Astron. Soc. **341**, 1333 (2003)
C. Adami, E. Slezak, F. Durret et al., Astron. Astrophys. **429**, 39 (2005)
A. Aguirre, L. Hernquist, J. Schaye et al., Astrophys. J. **561**, 521 (2001)
M. Arnaboldi, K.C. Freeman, S. Okamura et al., Astron. J. **125**, 514 (2003)
M. Arnaboldi, O. Gerhard, J.A.L. Aguerri et al., Astrophys. J. **614**, L33 (2004)
I. Balestra, P. Tozzi, P. Rosati et al., Astron. Astrophys. **462**, 429 (2007)
J. Barnes, L. Hernquist, Astron. Astrophys. **30**, 705 (1992)
J.H. Beall, J. Guillory, D.R. Rose, S. Schindler, S. Colafrancesco, Chin. J. Astron. Astrophys. Suppl. **3**, 137 (2004)
J.H. Beall, J. Guillory, D.R. Rose, S. Schindler, S. Colafrancesco, Chin. J. Astron. Astrophys. Suppl. **6**, 1 (2006)
K. Bekki, Astrophys. J. **510**, L15 (1999)
K. Bekki, W. Couch, Astrophys. J. **596**, L13 (2003)
G.M. Bernstein, R.C. Nichol, J.A. Tyson, M.P. Ulmer, D. Wittmann, Astron. J. **110**, 1507 (1995)
J. Binney, S. Tremaine, *Galactic Dynamics* (Princeton Univ. Press, Princeton, 1987)
E.L. Blanton, C.L. Sarazin, B.R. McNamara, M.W. Wise, Astrophys. J. **558**, L15 (2001)
S. Borgani, D. Fabjan, L. Tornatore et al., Space Sci. Rev. (2008). doi:10.1007/s11214-008-9322-7
H. Bravo-Alfaro, V. Cayatte, J.H. van Gorkom, C. Balkowski, Astron. J. **119**, 580 (2000)
H. Bravo-Alfaro, V. Cayatte, J.H. van Gorkom, C. Balkowski, Astron. Astrophys. **379**, 347 (2001)
D. Breitschwerdt, J.F. McKenzie, H.J. Voelk, Astron. Astrophys. **245**, 79 (1991)
M. Brüggen, C.R. Kaiser, Nature **418**, 301 (2002)

M. Brüggen, G. De Lucia, Mon. Not. R. Astron. Soc. **383**, 1336 (2008).
M. Busso, M. Marengo, C. Travaglio, L. Corcione, G. Silvestro, Mem. Soc. Astron. It. **72**, 309 (2001)
V. Cayatte, J.H. van Gorkom, C. Balkowski, C. Kotanyi, Astron. J. **100**, 604 (1990)
N. Caldwell, J.A. Rose, R.M. Sharples, R.S. Ellis, R.G. Bower, Astron. J. **106**, 473 (1993)
G. Cecil, J. Bland-Hawthorn, S. Veilleux, Astrophys. J. **576**, 745 (2002)
G. Chartas, W.N. Brandt, S.C. Gallagher, G.P. Garmire, Astrophys. J. **579**, 169 (2002)
G. Chartas, W.N. Brandt, S.C. Gallagher, Astrophys. J. **595**, 85 (2003)
Y. Choi, C.S. Reynolds, S. Heinz et al., Astrophys. J. **606**, 185 (2004)
E. Churazov, M. Brüggen, C.R. Kaiser, H. Böhringer, W. Forman, Astrophys. J. **554**, 261 (2001)
M.S. Clemens, P. Alexander, D.A. Green, Mon. Not. R. Astron. Soc. **312**, 236 (2000)
L. Cortese, G. Gavazzi, A. Boselli, J. Iglesias-Paramo, Astron. Astrophys. **416**, 119 (2004)
L. Cortese, D. Marcillac, J. Richard et al., Mon. Not. R. Astron. Soc. **376**, 157 (2007)
D.M. Crenshaw, S.B. Kraemer, I.M. George, Annu. Rev. Astron. Astrophys. **41**, 117 (2003)
H.H. Crowl, J.D.P. Kenney, J.H. van Gorkom, B. Vollmer, Astron. J. **130**, 65 (2005)
E.S. Cypriano, L. Sodre Jr., L.E. Campusano, D.A. Dale, E. Hardy, Astron. J. **131**, 2417 (2006)
S. Dado, A. Dar, A. De Rujula, Chin. J. Astron. Astrophys. (2007 in press). astro-ph/0703293
M. Dahlem, K.A. Weaver, T.M. Heckman, Astrophys. J. Suppl. Ser. **118**, 401 (1998)
C. Dalla Vecchia, R. Bower, T. Theuns et al., Mon. Not. R. Astron. Soc. **355**, 995 (2004)
S. Dasgupta, A.R. Rao, G.C. Dewangan, V.K. Agrawal, Astrophys. J. **618**, L87 (2005)
L.P. David, W. Forman, C. Jones, Astrophys. J. **380**, 39 (1991)
D.S. De Young, Astrophys. J. **223**, 47 (1978)
D.S. De Young, Astrophys. J. **307**, 62 (1986)
M.-H. Demoulin, E.M. Burbidge, Astrophys. J. **159**, 799 (1970)
W. Domainko, M. Gitti, S. Schindler, W. Kapferer, Astron. Astrophys. **425**, L21 (2004)
W. Domainko, M. Mair, W. Kapferer et al., Astron. Astrophys. **452**, 795 (2006)
A.E. Evrard, Mon. Not. R. Astron. Soc. **248**, 8 (1991)
A.C. Fabian, J.S. Sanders, G.B. Taylor et al., Mon. Not. R. Astron. Soc. **366**, 417 (2006)
J.J. Feldmeier, R. Ciardullo, G.H. Jacoby, Astrophys. J. **503**, 109 (1998)
J.J. Feldmeier, R. Ciardullo, G.H. Jacoby, P.R. Durrell, Astrophys. J. **615**, 196 (2004)
C. Ferrari, S. Maurogordato, A. Cappi, C. Benoist, Astron. Astrophys. **399**, 813 (2003)
C. Ferrari, C. Benoist, S. Maurogordato, A. Cappi, E. Slezak, Astron. Astrophys. **430**, 19 (2005)
C. Ferrari, R.W. Hunstead, L. Feretti, S. Maurogordato, S. Schindler, Astron. Astrophys. **457**, 21 (2006)
Y. Fujita, M. Takizawa, M. Nagashima, M. Enoki, Publ. Astron. Soc. Jpn. **51**, L1 (1999)
J.R. Gabel, N. Arav, T. Kim, Astrophys. J. **646**, 742 (2006)
A. Gal-Yam, D. Maoz, P. Guhathakurta, A.V. Filippenko, Astron. J. **125**, 1087 (2003)
R.D. Gehrz, J.W. Truran, R.E. Williams, S. Starrfield, Publ. Astron. Soc. Pac. **110**, 3 (1998)
O. Gerhard, M. Arnaboldi, K.C. Freeman, S. Okamura, Astrophys. J. **580**, L121 (2002)
O. Gerhard, M. Arnaboldi, K.C. Freeman et al., Astrophys. J. **621**, L93 (2005)
N.Y. Gnedin, Mon. Not. R. Astron. Soc. **294**, 407 (1998)
O.Y. Gnedin, Astrophys. J. **582**, 141 (2003)
A.H. Gonzalez, A.I. Zabludoff, D. Zaritsky, J.J. Dalcanton, Astrophys. J. **536**, 561 (2000)
A.H. Gonzalez, A.I. Zabludoff, D. Zaritsky, Astrophys. J. **618**, 195 (2005)
J.E. Gunn, J.R. Gott III, Astrophys. J. **176**, 1 (1972)
F.W. Hamann, T.A. Barlow, F.C. Chaffee, C.B. Foltz, R.J. Weymann, Astrophys. J. **550**, 142 (2001)
G. Hasinger, N. Schartel, S. Komossa, Astrophys. J. **573**, L77 (2002)
D. Heath, M. Krause, P. Alexander, Mon. Not. R. Astron. Soc. **374**, 787 (2007)
T.M. Heckman, M.D. Lehnert, D.K. Strickland, L. Armus, Astrophys. J. Suppl. Ser. **129**, 493 (2000)
T.M. Heckman, Rev. Mex. A **17**, 47 (2003)
S. Heinz, Y. Choi, C.S. Reynolds, M.C. Begelman, Astrophys. J. **569**, L79 (2002)
S. Heinz, E. Churazov, W. Forman, C. Jones, U.G. Briel, Mon. Not. R. Astron. Soc. **346**, 13 (2003)
S. Heinz, New Astron. Rev. **47**, 565 (2003)
S. Heinz, M. Brüggen, A. Young, E. Levesque, Mon. Not. R. Astron. Soc. **373**, L65 (2006)
J.A. Hester, Astrophys. J. **647**, 910 (2006)
M. Hidaka, Y. Sofue, Publ. Astron. Soc. Jpn. **54**, 33 (2002)
P. Jachym, J. Palous, J. Köppen, F. Combes, Astron. Astrophys. **472**, 5 (2007)
W. Kapferer, A. Knapp, S. Schindler, S. Kimeswenger, E. van Kampen, Astron. Astrophys. **438**, 87 (2005)
W. Kapferer, C. Ferrari, W. Domainko et al., Astron. Astrophys. **447**, 827 (2006)
W. Kapferer, T. Kronberger, J. Weratschnig et al., Astron. Astrophys. **466**, 813 (2007a)
W. Kapferer, T. Kronberger, J. Weratschnig, S. Schindler, Astron. Astrophys. **472**, 757 (2007b)
W. Kapferer, T. Kronberger, C. Ferrari, T. Riser, S. Schindler, Astron. Astrophys. (2008 submitted)
S.N. Kemp, E. de la Fuente, A. Franco-Balderas, J. Meaburn, Astrophys. J. **624**, 680 (2005)

J.D.P. Kenney, J.H. van Gorkom, B. Vollmer, Astron. J. **127**, 3361 (2004)
R.C. Kennicutt, Astrophys. J. **344**, 685 (1989)
R.C. Kennicutt, Astrophys. J. **498**, 541 (1998)
D.-W. Kim, E. Kim, G. Fabbiano, G. Trinchieri, Astrophys. J. (2007 submitted). astro-ph/0706.4254
R.A. Koopmann, J.D.P. Kenney, Astrophys. J. **613**, 866 (2004)
M. Krause, M. Camenzind, New Astron. Rev. **47**, 573 (2003)
J.E. Krick, R.A. Bernstein, Astron. J. **134**, 466 (2007)
T. Kronberger, W. Kapferer, C. Ferrari, S. Unterguggenberger, S. Schindler, Astron. Astrophys. (2008a submitted). astro-ph/0801.3759
T. Kronberger, W. Kapferer, D. Breitschwerdt, S. Schindler, in preparation (2008b)
L. Levy, J.A. Rose, J.H. van Gorkom, B. Chaboyer, Astron. J. **133**, 1104 (2007)
Y.-T. Lin, J.J. Mohr, Astrophys. J. **617**, 879 (2004)
D.M. Lucero, L.M. Young, J.H. van Gorkom, Astron. J. **129**, 647 (2005)
C.R. Lynds, A.R. Sandage, Astrophys. J. **137**, 1005 (1963)
M. Machacek, C. Jones, W.R. Forman, P. Nulsen, Astrophys. J. **644**, 155 (2006a)
M. Machacek, P.E.J. Nulsen, C. Jones, W. Forman, Astrophys. J. **648**, 947 (2006b)
A. Marcolini, F. Brighenti, A. D'Ercole, Mon. Not. R. Astron. Soc. **345**, 1329 (2003)
C.L. Martin, Astrophys. J. **513**, 156 (1999)
B.J. Maughan, C. Jones, W. Forman, L. Van Speybroeck, Astrophys. J. Suppl. Ser. **174**, 117 (2008)
L. Mayer, C. Mastropietro, J. Wadsley, J. Stadel, B. Moore, Mon. Not. R. Astron. Soc. **369**, 1021 (2006)
B.R. McNamara, M.W. Wise, P.E.J. Nulsen et al., Astrophys. J. **562**, L149 (2001)
B.R. McNamara, P.E.J. Nulsen, Annu. Rev. Astron. Astrophys. **45**, 117 (2007)
C.A. Metzler, A.E. Evrard, Astrophys. J. **437**, 564 (1994)
C.A. Metzler, A.E. Evrard, FERMILAB-Pub-97/349-A (1997). astro-ph/9710324
J.C. Mihos, P. Harding, J. Feldmeier, H. Morrison, Astrophys. J. **631**, L41 (2005)
B. Moore, N. Katz, G. Lake, A. Dressler, A. Oemler, Astrophys. J. **462**, 563 (1996)
R. Moll, S. Schindler, W. Domainko et al., Astron. Astrophys. **463**, 513 (2007)
M. Mori, A. Burkert, Astrophys. J. **538**, 559 (2000)
C. Moss, M. Whittle, Mon. Not. R. Astron. Soc. **317**, 667 (2000)
G. Murante, M. Giovalli, O. Gerhard et al., Mon. Not. R. Astron. Soc. **377**, 2 (2007)
N.P.H. Nesvadba, M.D. Lehnert, F. Eisenhauer et al., Astrophys. J. **650**, 693 (2006)
B. Nordström, Astrophys. Space Sci. **284**, 779 (2003)
P.E.J. Nulsen, L.P. David, B.R. McNamara et al., Astrophys. J. **568**, 163 (2002)
P.T. O'Brian, J.N. Reeves, C. Simpson, M.J. Ward, Mon. Not. R. Astron. Soc. **360**, L25 (2005)
T. Oosterloo, J. van Gorkom, Astron. Astrophys. **437**, L19 (2005)
K. Otmianowska-Mazur, B. Vollmer, Astron. Astrophys. **402**, 879 (2003)
F.N. Owen, M.J. Ledlow, W.C. Keel, G.E. Morrison, Astron. J. **118**, 633 (1999)
K.A. Pounds, J.N. Reeves, A.R. King et al., Mon. Not. R. Astron. Soc. **345**, 705 (2003a)
K.A. Pounds, A.R. King, K.L. Page, P.T. O'Brien, Mon. Not. R. Astron. Soc. **346**, 1025 (2003b)
V. Quilis, B. Moore, R. Bower, Science **288**, 1617 (2000)
F.V.N. Rangarajan, D.A. White, H. Ebeling, A.C. Fabian, Mon. Not. R. Astron. Soc. **277**, 1047 (1995)
J. Rasmussen, T.J. Ponman, J.S. Mulchaey, Mon. Not. R. Astron. Soc. **370**, 453 (2006)
E. Rasia, P. Mazzotta, H. Bourdin et al., Astrophys. J. (2007 in press). astro-ph/0707.2614
J.N. Reeves, P.T. O'Brien, M.J. Ward, Astrophys. J. **593**, L65 (2003)
E. Roediger, G. Hensler, Astron. Astrophys. **433**, 875 (2005)
E. Roediger, M. Brüggen, Mon. Not. R. Astron. Soc. **369**, 567 (2006)
E. Roediger, M. Brüggen, M. Hoeft, Mon. Not. R. Astron. Soc. **371**, 609 (2006)
E. Roediger, M. Brüggen, Mon. Not. R. Astron. Soc. **380**, 1399 (2007)
E. Roediger, M. Brüggen, P. Rebusco, H. Böhringer, E. Churazov, Mon. Not. R. Astron. Soc. **375**, 15 (2007)
C.S. Rudick, J.C. Mihos, C. McBride, Astrophys. J. **648**, 936 (2006)
D.S. Rupke, S. Veilleux, D.B. Sanders, Astrophys. J. **570**, 588 (2002)
E.V. Ryan-Weber, G.R. Meurer, K.C. Freeman et al., Astron. J. **127**, 1431 (2004)
S. Schindler, A. Castillo-Morales, E. De Filippis, A. Schwope, J. Wambsganß, Astron. Astrophys. **376**, L27 (2001)
S. Schindler, W. Kapferer, W. Domainko et al., Astron. Astrophys. **435**, L25 (2005)
M. Schmidt, Astrophys. J. **129**, 243 (1959)
S. Schulz, C. Struck, Mon. Not. R. Astron. Soc. **328**, 185 (2001)
D. Sijacki, V. Springel, Mon. Not. R. Astron. Soc. **366**, 397 (2006)
V. Springel, L. Hernquist, Mon. Not. R. Astron. Soc. **339**, 289 (2003)
V. Springel, T. Di Matteo, L. Hernquist, Mon. Not. R. Astron. Soc. **361**, 776 (2005)
D.K. Strickland, I.R. Stevens, Mon. Not. R. Astron. Soc. **314**, 511 (2000)

D.K. Strickland, T.M. Heckman, K.A. Weaver, C.G. Hoopes, M. Dahlem, Astrophys. J. **568**, 689 (2002)
M. Sun, M. Donahue, M. Voit, Astrophys. J. **671**, 190 (2007)
G. Tenorio-Tagle, C. Munoz-Tunon, Mon. Not. R. Astron. Soc. **293**, 299 (1998)
T. Theuns, S.J. Warren, Mon. Not. R. Astron. Soc. **284**, L11 (1997)
T. Toniazzo, S. Schindler, Mon. Not. R. Astron. Soc. **325**, 509 (2001)
L.B. van den Hoek, M.A.T. Groenewegen, Astron. Astrophys. Suppl. Ser. **123**, 305 (1997)
J. van Gorkom, in *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, ed. by J.S. Mulchaey, A. Dressler, A. Oemler (Cambridge Univ. Press, Cambridge, 2003). astro-ph/0308209
S. Veilleux, J. Bland-Hawthorn, G. Cecil, R.B. Tully, S.T. Miller, Astrophys. J. **520**, 111 (1999)
S. Veilleux, G. Cecil, J. Bland-Hawthorn, Annu. Rev. Astron. Astrophys. **43**, 769 (2005)
B. Vollmer, V. Cayatte, A. Boselli, C. Balkowski, W.J. Duschl, Astron. Astrophys. **349**, 411 (1999)
B. Vollmer, V. Cayatte, C. Balkowski, W.J. Duschl, Astrophys. J. **561**, 708 (2001)
B. Vollmer, Astron. Astrophys. **398**, 525 (2003)
B. Vollmer, C. Balkowski, V. Cayatte, W. van Driel, W. Huchtmeier, Astron. Astrophys. **419**, 35 (2004a)
B. Vollmer, R. Beck, J.D.P. Kenney, J.H. van Gorkom, Astron. J. **127**, 3375 (2004b)
B. Vollmer, M. Soida, K. Otmianowska-Mazur et al., Astron. Astrophys. **453**, 883 (2006)
B. Vollmer, W. Huchtmeier, Astron. Astrophys. **462**, 93 (2007)
Q.D. Wang, M.P. Ulmer, R.J. Lavery, Mon. Not. R. Astron. Soc. **288**, 702 (1997)
Q.D. Wang, S. Immler, R. Walterbos, J.T. Lauroesch, D. Breitschwerdt, Astrophys. J. **555**, L99 (2001)
N. Werner, F. Durret, T. Ohashi, S. Schindler, R.P.C. Wiersma, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9320-9
B.F. Williams, R. Ciardullo, P.R. Durrell et al., Astrophys. J. **656**, 756 (2007)
B. Willman, F. Governato, J. Wadsley, T. Quinn, Mon. Not. R. Astron. Soc. **355**, 159 (2004)
J.C. Wheeler, C. Sneden, J.W. Truran, Annu. Rev. Astron. Astrophys. **27**, 279 (1989)
C. Zanni, G. Murante, G. Bodo et al., Astron. Astrophys. **429**, 399 (2005)
D. Zaritzsky, A.H. Gonzales, A.I. Zabludoff, Astrophys. J. **613**, L93 (2004)
H. Zhang, S. Koide, J. Sakai, Publ. Astron. Soc. Jpn. **51**, 449 (1999)
S. Zibetti, S.D.M. White, D.P. Schneider, J. Brinkmann, Mon. Not. R. Astron. Soc. **358**, 949 (2005)
A. Zijlstra, in *Proc. of IAU Symp. 234, Planetary Nebulae in the Galaxy and Beyond*, ed. by M.J. Barlow, R.H. Mendez (Cambridge Univ. Press, Cambridge, 2006), p. 55. astro-ph/0605097

Chapter 18
The Chemical Enrichment of the ICM from Hydrodynamical Simulations

S. Borgani · D. Fabjan · L. Tornatore · S. Schindler · K. Dolag · A. Diaferio

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9322-7

Abstract The study of the metal enrichment of the intra-cluster and inter-galactic media (ICM and IGM) represents a direct means to reconstruct the past history of star formation, the role of feedback processes and the gas-dynamical processes which determine the evolution of the cosmic baryons. In this paper we review the approaches that have been followed so far to model the enrichment of the ICM in a cosmological context. While our presentation will be focused on the role played by hydrodynamical simulations, we will also discuss

S. Borgani (✉) · D. Fabjan · L. Tornatore
Department of Astronomy, University of Trieste, via Tiepolo 11, 34143 Trieste, Italy
e-mail: borgani@oats.inaf.it

D. Fabjan
e-mail: fabjan@oats.inaf.it

L. Tornatore
e-mail: tornatore@oats.inaf.it

S. Borgani · D. Fabjan
INAF—National Institute for Astrophysics, Trieste, Italy

S. Borgani · D. Fabjan
INFN—National Institute for Nuclear Physics, Sezione di Trieste, Trieste, Italy

S. Schindler
Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria
e-mail: Sabine.Schindler@uibk.ac.at

K. Dolag
Max-Planck-Institut für Astrophysik, Karl-Schwarzschild Strasse 1, Garching bei München, Germany
e-mail: kdolag@mpa-garching.mpg.de

A. Diaferio
Dipartimento di Fisica Generale "Amedeo Avogadro", Università degli Studi di Torino, Torino, Italy
e-mail: diaferio@ph.unito.it

A. Diaferio
INFN—National Institute for Nuclear Physics, Sezione di Torino, Torino, Italy

other approaches based on semi-analytical models of galaxy formation, also critically discussing pros and cons of the different methods. We will first review the concept of the model of chemical evolution to be implemented in any chemo-dynamical description. We will emphasise how the predictions of this model critically depend on the choice of the stellar initial mass function, on the stellar life-times and on the stellar yields. We will then overview the comparisons presented so far between X-ray observations of the ICM enrichment and model predictions. We will show how the most recent chemo-dynamical models are able to capture the basic features of the observed metal content of the ICM and its evolution. We will conclude by highlighting the open questions in this study and the direction of improvements for cosmological chemo-dynamical models of the next generation.

Keywords Cosmology: numerical simulations · Galaxies: clusters · Hydrodynamics · X-ray: galaxies

1 Introduction

Clusters of galaxies are the ideal cosmological signposts to trace the past history of the intergalactic medium (IGM), thanks to the high density and temperature reached by the cosmic baryons trapped in their gravitational potential wells (Rosati et al. 2002; Voit 2005; Diaferio et al. 2008—Chap. 2, this issue). Observations in the X-ray band with the Chandra and XMM-Newton satellites are providing invaluable information on the thermodynamical properties of the intra-cluster medium (ICM) (Kaastra et al. 2008—Chap. 9, this issue). These observations highlight that non-gravitational sources of energy, such as energy feedback from supernovae (SNe) and Active Galactic Nuclei (AGN) have played an important role in determining the ICM physical properties. Spatially resolved X-ray spectroscopy permits to measure the equivalent width of emission lines associated to transitions of heavily ionised elements and, therefore, to trace the pattern of chemical enrichment (e.g., Mushotzky 2004 for a review). In turn, this information is inextricably linked to the history of formation and evolution of the galaxy population (e.g., Renzini 1997 and references therein), as inferred from observations in the optical/near-IR band. For instance, De Grandi et al. (2004) have first shown that cool core clusters are characterised by a significant central enhancement of the iron abundance, which closely correlates with the magnitude of the Brightest Cluster Galaxies (BCGs). This demonstrates that a fully coherent description of the evolution of cosmic baryons in the condensed stellar phase and in the diffuse hot phase requires properly accounting for the mechanisms of production and release of both energy and metals.

The study of how these processes take place during the hierarchical build up of cosmic structures has been tackled with different approaches. Semi-analytical models (SAMs) of galaxy formation provide a flexible tool to explore the space of parameters which describe a number of dynamical and astrophysical processes. In their most recent formulation, such models are coupled to dark matter (DM) cosmological simulations, to trace the merging history of the halos where galaxy formation takes place, and include a treatment of metal production from type-Ia and type-II supernovae (SN Ia and SN II, hereafter; De Lucia et al. 2004; Nagashima et al. 2005), so as to properly address the study of the chemical enrichment of the ICM. The main limitation of this method is that it does not include the gas dynamical processes which causes metals, once produced, to be transported in the ICM. As a consequence, they provide a description of the global metallicity of clusters and its evolution, but not of the details of its spatial distribution.

In order to overcome this limitation, Cora (2006) applied an alternative approach, in which non-radiative hydrodynamical simulations of galaxy clusters are used to trace at the

same time the formation history of DM halos and the dynamics of the gas. In this approach, metals are produced by SAM galaxies and then suitably assigned to gas particles, thereby providing a chemo-dynamical description of the ICM. Domainko et al. (2006) used hydrodynamical simulations, which include prescriptions for gas cooling, star formation and feedback, similar to those applied in SAMs, to address the specific role played by ram-pressure stripping to distribute metals, while Kapferer et al. (2007b) used the same approach to study the different roles played by galactic winds and by ram-pressure stripping. While these approaches offer advantages with respect to standard SAMs, they still do not provide a fully self-consistent picture, in which chemical enrichment is the outcome of the process of star formation, associated to the cooling of the gas infalling in high density regions. In this sense, a fully self-consistent approach requires that the simulations include the processes of gas cooling, star formation and evolution, along with the corresponding feedback in energy and metals.

A number of authors have presented hydrodynamical simulations for the formation of cosmic structures, which include treatments of the chemical evolution at different levels of complexity. Raiteri et al. (1996) presented SPH simulations of the Galaxy, forming in an isolated halo, by following iron and oxygen production from SN II and SN Ia, also accounting for the effect of stellar lifetimes. Mosconi et al. (2001) performed a detailed analysis of chemo-dynamical SPH simulations, aimed at studying both numerical stability of the results and the enrichment properties of galactic objects in a cosmological context. Lia et al. (2002) discussed a statistical approach to follow metal production in SPH simulations, which have a large number of star particles, showing applications to simulations of a disc-like galaxy and of a galaxy cluster. Kawata and Gibson (2003b) carried out cosmological chemo-dynamical simulations of elliptical galaxies, with an SPH code, by including the contribution from SN Ia and SN II, also accounting for stellar lifetimes. Valdarnini (2003) performed an extended set of cluster simulations and showed that profiles of the iron abundance are steeper than the observed ones. A similar analysis has been presented by Romeo et al. (2006), who also considered the effect of varying the IMF and the feedback efficiency on the enrichment pattern of the ICM. Scannapieco et al. (2005) presented an implementation of a model of chemical enrichment in the `GADGET-2` code, coupled to a self-consistent model for star formation and feedback. In their model, which was applied to study the enrichment of galaxies, they included the contribution from SN Ia and SN II, assuming that all long-lived stars die after a fixed delay time. Tornatore et al. (2004, 2007a) presented results from an implementation of a detailed model of chemical evolution model in the `GADGET-2` code (Springel 2005), including metallicity-dependent yields and the contribution from intermediate and low mass stars (ILMS hereafter). The major advantage of this approach is that the metal production is self-consistently predicted from the rate of gas cooling treated by the hydrodynamical simulations, without resorting to any external model. However, at present the physical scales involved by the processes of star formation and SN explosions are far from being resolved in simulations which sample cosmological scales. For this reason, such simulations also need to resort to external sub-resolution models, which provide an effective description of a number of relevant astrophysical processes.

The aim of this paper is to provide a review of the results obtained so far in the study of the chemical enrichment of the ICM in a cosmological context. Although we will concentrate the discussion on the results obtained from full hydrodynamical simulations, we will also present results based on SAMs. As such, this paper complements the reviews by Borgani et al. (2008)—Chap. 13, this issue, which reviews the current status in the study of the thermodynamical properties of the ICM with cosmological hydrodynamical simulations, and by Schindler and Diaferio (2008)—Chap. 17, this issue, which will focus on the study of

the role played by different physical processes in determining the ICM enrichment pattern. Also, this paper will not present a detailed description of the techniques for simulations of galaxy clusters, which is reviewed by Dolag et al. (2008)—Chap. 12, this issue.

In Sect. 2 we review the concept of model of chemical evolution and highlight the main quantities which are required to fully specify this model. In Sect. 3 we review the results on the global metal content of the ICM, while Sect. 4 concentrates on the study of the metallicity profiles and Sect. 5 on the study of the ICM evolution. Section 6 discusses the properties of the galaxy population. Finally, Sect. 7 provides a critical summary of the results presented, by highlighting the open problems and lines of developments to be followed by simulations of the next generation.

2 What is a model of chemical evolution?

In this section we provide a basic description of the key ingredients required by a model of chemical evolution. For a more detailed description we refer to the book by Matteucci (2003).

The process of star formation in cosmological hydrodynamical simulations is described through the conversion of a gas element into a star particle; as a consequence, in simulations of large volumes the star particles have a mass far larger than that of a single star, with values of the order of 10^6–10^8 $M_\odot$, depending on the resolution and on the mass of the simulated structure (e.g., Katz et al. 1996). The consequence of this coarse-grained representation of star formation is that each star particle must be treated as a simple stellar population (SSP), i.e. as an ensemble of coeval stars having the same initial metallicity. Every star particle carries all the physical information (e.g. birth time t_b, initial metallicity and mass) that is needed to calculate the evolution of the stellar populations that they represent, once the lifetime function (see Sect. 2.2), the IMF (see Sect. 2.4) and the yields (see Sect. 2.3) for SNe and ILMS have been specified. Therefore, we can compute for every star particle at any given time $t > t_b$ how many stars are dying as SN II and SN Ia, and how many stars undergo the AGB phase, according to the equations of chemical evolution, that we will present here below.

The generation of star particles is pertinent to hydrodynamical simulations, which include the description of the star formation. As for semi-analytic models (SAMs) of galaxy formation, they generally identify a galaxy with a suitably chosen DM particle in the collisionless simulation which is used to reconstruct the merger tree. While this DM particle defines the position of a galaxy, the corresponding stellar content is described as a superposition of SSPs, each generated in correspondence of the time step with which the process of galaxy formation is studied (e.g., Diaferio et al. 2001; De Lucia et al. 2004; Nagashima et al. 2005, for applications of SAMs to the cluster galaxies). In this sense, the description of the chemical evolution model, as we provide here, can be used for both hydrodynamical simulations and for SAMs.

It is generally assumed that the stars having mass above 8 $M_\odot$ at the end of the hydrostatic core burning undergo an electron capture process, leading to a core collapse. A large amount of energy can be transferred to the outer layers during this phase due to a substantial production of neutrinos that easily escape from the central core. Although theoretical work has not yet been able to reproduce a sufficient energy deposition, it is currently supposed that this process leads to an explosive ejection of the outer layers, giving rise to a SN II. We remind the reader that 8 $M_\odot$ is a commonly adopted fiducial value, although the limiting mass for the onset of explosive evolution is still debated (e.g., Portinari et al. 1998).

A different ejection channel is provided by the SN Ia that are believed to arise from thermonuclear explosions of white dwarfs (WD hereafter) in binary stellar systems as a consequence of the matter accretion from the companion (e.g., Nomoto et al. 2000). However, there are still a number of uncertainties about the nature of both the WD and the companion and about the mass reached at the onset of the explosion (e.g., Matteucci and Recchi 2001; Yungelson and Livio 2000). Furthermore, observational evidences for multiple populations of Type Ia progenitors have been recently found both from direct detection of Type Ia events (e.g., Mannucci et al. 2005; Scannapieco and Bildsten 2005; Sullivan et al. 2006) or from inferences on the chemical enrichment patterns in galaxy clusters (e.g., Gal-Yam and Maoz 2004). Finally, a third way to eject heavy elements in the interstellar medium is the mass loss of ILMS by stellar winds.

In summary, the main ingredients that define a model of chemical evolution are the following: (a) the SNe explosion rates, (b) the adopted lifetime function, (c) the adopted yields and (d) the IMF which fixes the number of stars of a given mass. We describe each of these ingredients in the following.

2.1 The Equations of Chemical Evolution

2.1.1 Type Ia Supernovae

We provide here below a short description of how Type Ia SN are included in a model of chemical evolution. For a comprehensive review of analytical formulations we refer to the paper by Greggio (2005). Following Greggio and Renzini (1983), we assume here that SN Ia arise from stars belonging to binary systems, having mass in the range 0.8–8 $M_{\odot}$. Accordingly, in the single-degenerate WD scenario (Nomoto et al. 2000), the rate of explosions of SN Ia can be written as

$$R_{\mathrm{SN\,Ia}}(t) = A \int_{M_{\mathrm{B,inf}}}^{M_{\mathrm{B,sup}}} \phi(m_{\mathrm{B}}) \int_{\mu_{\mathrm{m}}}^{\mu_{\mathrm{M}}} f(\mu)\, \psi(t - \tau_{m_2})\, \mathrm{d}\mu\, \mathrm{d}m_{\mathrm{B}}. \quad (1)$$

In the above equation, $\phi(m)$ is the IMF, m_{B} is the total mass of the binary system, m_2 is the mass of the secondary companion, τ_m is the mass-dependent life-time and $\psi(t)$ is the star formation rate. The variable $\mu = m_2/m_{\mathrm{B}}$ is distributed according to the function $f(\mu)$, while A is the fraction of stars in binary systems of that particular type to be progenitors of SN Ia (see Matteucci and Recchi 2001 for more details). For instance, in the model by Greggio and Renzini (1983) μ varies in the range between μ_{m} and $\mu_{\mathrm{M}} = 0.5$, with $\mu_{\mathrm{m}} = \max[m_2(t)/m_{\mathrm{B}}, (m_{\mathrm{B}} - 0.5 M_{\mathrm{BM}})/m_{\mathrm{B}}]$, where $m_2(t)$ is the mass of the companion which dies at the time t, according to the chosen life-time function. Furthermore, let M_{Bm} and M_{BM} be the smallest and largest value allowed for the progenitor binary mass m_{B}. Then, the integral over m_{B} runs in the range between $M_{\mathrm{B,inf}}$ and $M_{\mathrm{B,sup}}$, which represent the minimum and the maximum value of the total mass of the binary system that is allowed to explode at the time t. These values in general are functions of M_{Bm}, M_{BM}, and $m_2(t)$, which in turn depends on the star formation history $\Psi(t)$. The exact functional dependence is defined by the SN Ia progenitor model. For instance, in the model by Greggio and Renzini (1983) it is $M_{\mathrm{B,inf}} = \max[2m_2(t), M_{\mathrm{Bm}}]$ and $M_{\mathrm{B,sup}} = 0.5 M_{\mathrm{BM}} + m_2(t)$. Under the assumption of a short duration burst of star formation, the function $\psi(t)$ can be approximated with a Dirac δ-function. This case applies to hydrodynamical simulations, which include star formation, where the creation of a SSP is described by an impulsive star formation event, while more complex descriptions should take into account the continuous star formation history $\psi(t)$. Under the

above assumption for $\psi(t)$ and using the functional form of $f(\mu)$ derived from statistical studies of the stellar population in the solar neighbourhood (Tutukov and Iungelson 1980; Matteucci and Recchi 2001), we find

$$R_{\rm SN\,Ia}(t) = -\frac{{\rm d}m_2(t)}{{\rm d}t}\bigg|_{m_2\equiv\tau^{-1}(t)} 24 m_2^2 A \int_{M_{\rm Bm}}^{M_{\rm BM}} \phi(m_{\rm B})\frac{1}{m_{\rm B}^3}{\rm d}m_{\rm B}. \quad (2)$$

Since the current understanding of the process of star formation does not allow to precisely determine the value of A, its choice can be fixed from the requirement of reproducing a specific observation, once the form of the IMF is fixed. For instance, Matteucci and Gibson (1995) found that $A = 0.1$ was required to reproduce the observed iron enrichment of the ICM.

2.1.2 Supernova Type II and Low and Intermediate Mass Stars

Computing the rates of SN II and ILMS is conceptually simpler, since they are driven by the lifetime function $\tau(m)$ convolved with the star formation history $\psi(t)$ and multiplied by the IMF $\phi(m = \tau^{-1}(t))$. Again, since $\psi(t)$ is a delta-function for the SSP used in simulations, the SN II and ILMS rates read

$$R_{\rm SN\,II|ILMS}(t) = \phi(m(t)) \times \left(-\frac{{\rm d}m(t)}{{\rm d}t}\right) \quad (3)$$

where $m(t)$ is the mass of the star that dies at time t. We note that the above expression must be multiplied by a factor of $(1 - A)$ for AGB rates if the interested mass $m(t)$ falls in the same range of masses which is relevant for the secondary stars of SN Ia binary systems.

In order to compute the metal release by stars (binary systems in case of SN Ia) of a given mass we need to take into account the yields $p_{Z_i}(m, Z)$, which provide the mass of the element i produced by a star of mass m and initial metallicity Z. Then, the equation which describes the evolution of the mass $\rho_i(t)$ for the element i, holding for a generic form of the star formation history $\psi(t)$, reads:

$$\begin{aligned}\dot{\rho}_i(t) = &-\psi(t)Z_i(t)\\ &+ A\int_{M_{\rm Bm}}^{M_{\rm BM}} \phi(m)\left[\int_{\mu_{\rm min}}^{0.5} f(\mu)\psi(t-\tau_{m_2})p_{Z_i}(m,Z){\rm d}\mu\right]{\rm d}m\\ &+ (1-A)\int_{M_{\rm Bm}}^{M_{\rm BM}} \psi(t-\tau(m))p_{Z_i}(m,Z)\varphi(m){\rm d}m\\ &+ \int_{M_{\rm L}}^{M_{\rm Bm}} \psi(t-\tau(m))p_{Z_i}(m,Z)\varphi(m){\rm d}m\\ &+ \int_{M_{\rm BM}}^{M_{\rm U}} \psi(t-\tau(m))p_{Z_i}(m,Z)\varphi(m){\rm d}m. \end{aligned} \quad (4)$$

In the above equation, $M_{\rm L}$ and $M_{\rm U}$ are the minimum and maximum mass of a star, respectively. Commonly adopted choices for these limiting masses are $M_{\rm L} \simeq 0.1$ M$_\odot$ and $M_{\rm U} \simeq 100$ M$_\odot$. The term in the first line of (4) accounts for the metallicity sink due to the locking of metals in the newborn stars. The term in the second line accounts for metal ejection contributed by SN Ia. The terms in the third and fourth lines describe the enrichment by mass-loss from intermediate and low mass stars, while the last line accounts for ejecta by SN II.

2.2 The Lifetime Function

Different choices for the mass-dependence of the life-time function have been proposed in the literature. For instance, Padovani and Matteucci (1993) (PM93 hereafter) proposed the expression

$$\tau(m) = \begin{cases} 10^{[(1.34-\sqrt{1.79-0.22(7.76-\log(m))})/0.11]-9} & \text{for } m \leq 6.6\ \mathrm{M}_\odot, \\ 1.2m^{-1.85} + 0.003 & \text{otherwise.} \end{cases} \quad (5)$$

An alternative expression has been proposed by Maeder and Meynet (1989) (MM89 hereafter), and extrapolated by Chiappini et al. (1997) to very high ($> 60\ \mathrm{M}_\odot$) and very low ($< 1.3\ \mathrm{M}_\odot$) masses:

$$\tau(m) = \begin{cases} 10^{-0.6545\log m+1} & m \leq 1.3\ \mathrm{M}_\odot, \\ 10^{-3.7\log m+1.351} & 1.3 < m \leq 3\ \mathrm{M}_\odot, \\ 10^{-2.51\log m+0.77} & 3 < m \leq 7\ \mathrm{M}_\odot, \\ 10^{-1.78\log m+0.17} & 7 < m \leq 15\ \mathrm{M}_\odot, \\ 10^{-0.86\log m-0.94} & 15 < m \leq 53\ \mathrm{M}_\odot, \\ 1.2 \times m^{-1.85} + 0.003 & \text{otherwise.} \end{cases} \quad (6)$$

The main difference between these two functions concerns the life-time of low mass stars ($< 8\mathrm{M}_\odot$). The MM89 life-time function delays the explosion of stars with mass $\gtrsim 1\ \mathrm{M}_\odot$, while it anticipates the explosion of stars below $1\ \mathrm{M}_\odot$ with respect to the PM93 life-time function. Only for masses below $1\ \mathrm{M}_\odot$ does the PM93 function predict much more long-living stars. This implies that different life-times will produce different evolution of both absolute and relative abundances (we refer to Romano et al. 2005 for a detailed description of the effect of the lifetime function in models of chemical evolution).

We point out that the above lifetime functions are independent of metallicity, whereas in principle this dependence can be included in a model of chemical evolution (e.g., Portinari et al. 1998). For instance, Raiteri et al. (1996) used the metallicity-dependent lifetimes as obtained from the Padova evolutionary tracks (Bertelli et al. 1994). We show in Fig. 1 a comparison of different lifetime functions.

2.3 Stellar Yields

The stellar yields specify the quantity $p_{Z_i}(m, Z)$, which appears in (4) and, therefore, the amount of different metal species which are released during the evolution of a SSP. A number of different sets of yields have been proposed in the literature, such as those by Renzini and Voli (1981), van den Hoek and Groenewegen (1997), Marigo (2001) for the LIMS and those by Nomoto et al. (1997), Iwamoto et al. (1999), Thielemann et al. (2003) for SN Ia. As for SN II, there are many proposed sets of metallicity-dependent yields; among others, those by Woosley and Weaver (1995), by Portinari et al. (1998), by Chieffi and Limongi (2004), which are based on different assumptions of the underlying model of stellar structure and evolution.

As an example of the differences among different sets of yields, we show in the left panel of Fig. 2 the ratios between the abundances of different elements produced by the SN II of a SSP, as expected from Woosley and Weaver (1995) and from Chieffi and Limongi (2004). Different curves and symbols here correspond to different values of the initial metallicity of

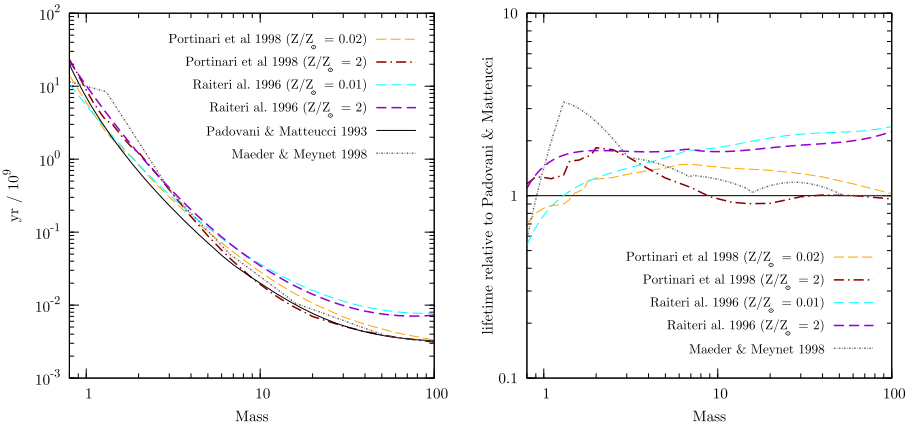


Fig. 1 *Left panel*: the mass dependence of different lifetime functions: Padovani and Matteucci (1993, *solid line*), Maeder and Meynet (1989, *dotted line*), Raiteri et al. (1996, *dashed lines*; the *thin line* refers to $Z = 0.02Z_{\odot}$, the *thick line* refers to $Z = 2Z_{\odot}$), Portinari et al. (1998, *dot-dashed lines*; the *thin line* refers to $Z = 0.02Z_{\odot}$, the *thick line* refers to $Z = 2Z_{\odot}$). *Right panel*: the ratio between the same lifetimes and that by Padovani and Matteucci (1993)

the SSP. Quite apparently, the two sets of yields provide significantly different metal masses, by an amount which can sensitively change with initial metallicity. This illustrates how a substantial uncertainty exists nowadays about the amount of metals produced by different stellar populations. There is no doubt that these differences between different sets of yields represent one of the main uncertainties in any modelling of the chemical evolution of the ICM.

For this reason, we recommend that papers related to the cosmological modelling of the chemical evolution, either by SAMs or by hydrodynamical simulations, should be very careful in specifying for which set of yields the computations have been performed as well as the details of assumed lifetime and models for SN II and SN Ia. In the absence of this, it becomes quite hard to judge the reliability of any detailed comparison with observational data or with the predictions of other models.

2.4 The Initial Mass Function

The initial mass function (IMF) is one of the most important quantities in a model of chemical evolution. It directly determines the relative ratio between SN II and SN Ia and, therefore, the relative abundance of α-elements and Fe-peak elements. The shape of the IMF also determines how many long-living stars will form with respect to massive short-living stars. In turn, this ratio affects the amount of energy released by SNe and the present luminosity of galaxies, which is dominated by low mass stars, and the (metal) mass-locking in the stellar phase.

As of today, no general consensus has been reached on whether the IMF at a given time is universal or strongly dependent on the environment, or whether it is time-dependent, i.e. whether local variations of the values of temperature, pressure and metallicity in star-forming regions affect the mass distribution of stars.

The IMF $\phi(m)$ is defined as the number of stars of a given mass per unit logarithmic mass interval. A widely used form is

$$\phi(m) = dN/d\log m \propto m^{-x(m)}. \tag{7}$$

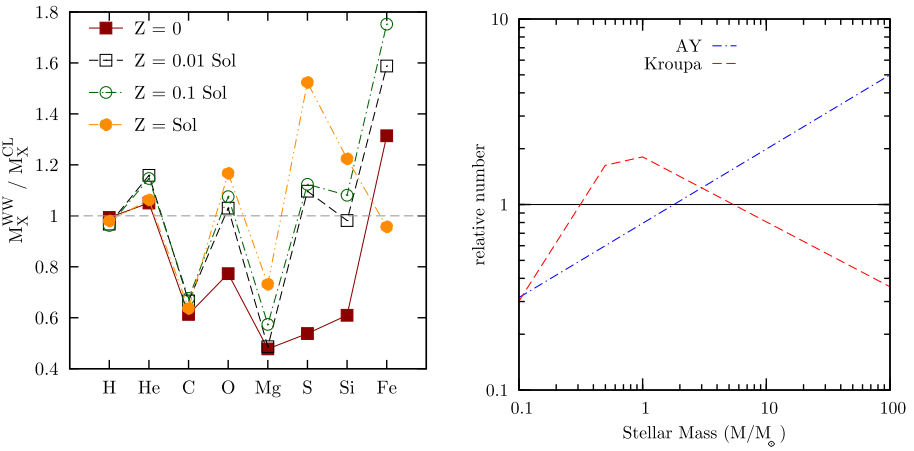


Fig. 2 *Left panel*: the ratio $M_j^{\rm WW}/M_j^{\rm CL}$ between the mass of species j, produced by the SN II of a SSP, when using the two sets of yields by Woosley and Weaver (1995) and by Chieffi and Limongi (2004) for different values of the initial SSP metallicity. Different symbols are for different values of the initial metallicity of the SSP, as reported by the labels. From Tornatore et al. (2007a). *Right panel*: the shape of different IMFs introduced in the literature. Stellar masses are expressed in units of $M_\odot$

If the exponent x in the above expression does not depend on the mass m, the IMF is then described by a single power-law. The most famous and widely used single power-law IMF is the Salpeter (1955) one that has $x = 1.35$. Arimoto and Yoshii (1987) proposed an IMF with $x = 0.95$, which predicts a relatively larger number of massive stars. In general, IMFs providing a large number of massive stars are usually called top-heavy. More recently, different expressions of the IMF have been proposed in order to model a flattening in the low-mass regime that is currently favoured by a number of observations. Kroupa (2001) introduced a multi-slope IMF, which is defined as

$$\phi(m) \propto \begin{cases} m^{-1.3} & m \geq 0.5 M_\odot, \\ m^{-0.3} & 0.08 \leq m < 0.5 M_\odot, \\ m^{\,0.7} & m \leq 0.08 M_\odot. \end{cases} \tag{8}$$

Chabrier (2003) proposed another expression for the IMF, which is quite similar to that one proposed by Kroupa

$$\phi(m) \propto \begin{cases} m^{-1.3} & m > 1 M_\odot, \\ e^{\frac{-(\log(m)-\log(m_c))^2}{2\sigma^2}} & m \leq 1 M_\odot. \end{cases} \tag{9}$$

Theoretical arguments (e.g. Larson 1998) suggest that the present-day characteristic mass scale, where the IMF changes its slope, $\sim 1\ M_\odot$ should have been larger in the past, so that the IMF at higher redshift was top-heavier than at present.

While the shape of the IMF is determined by the local conditions of the inter-stellar medium, direct hydrodynamical simulations of star formation in molecular clouds are only now approaching the required resolution and sophistication level to make credible predictions on the IMF (e.g., Bonnell et al. 2006, Padoan et al. 2007; see McKee and Ostriker 2007 for a detailed discussion).

We show in Fig. 2 the number of stars, as a function of their mass, predicted by different IMFs, relative to those of the Salpeter IMF. As expected, the AY IMF predicts a larger number of high-mass stars and, correspondingly, a smaller number of low-mass stars, the crossover taking place at $\simeq 2\ M_\odot$. As a result, we expect that the enrichment pattern of the AY IMF will be characterised by a higher abundance of those elements, like oxygen, which are mostly produced by SN II. Both the Kroupa and the Chabrier IMFs are characterised by a relative deficit of very low-mass stars and a mild overabundance of massive stars and ILMS. Correspondingly, an enhanced enrichment in both Fe-peak and α elements like oxygen is expected, mostly due to the lower fraction of mass locked in ever-living stars. For reference we also show a different IMF, also proposed by Kroupa et al. (1993), that exhibits a deficit in both very low- and high-mass stars; for this kind of IMF a lower α/Fe ratio with respect to the Salpeter IMF is expected.

Since clusters of galaxies basically behave like "closed boxes", the overall level of enrichment and the relative abundances should directly reflect the condition of star formation. While a number of studies have been performed so far to infer the IMF shape from the enrichment pattern of the ICM, no general consensus has been reached. For instance, Renzini (1997, 2004) and Pipino et al. (2002) argued that both the global level of ICM enrichment and the α/Fe relative abundance can be reproduced by assuming a Salpeter IMF, as long as this relative abundance is sub-solar. Indeed, a different conclusion has been reached by other authors in the attempt of explaining the relative abundances $[\alpha/\mathrm{Fe}] \gtrsim 0$ in the ICM (e.g., Baumgartner et al. 2005). Portinari et al. (2004) used a phenomenological model to argue that a standard Salpeter IMF cannot account for the observed α/Fe ratio in the ICM. As we shall discuss in the following, a similar conclusion was also reached by Nagashima et al. (2005), who used semi-analytical models of galaxy formation to trace the production of heavy elements, and by Romeo et al. (2005), who used hydrodynamical simulations including chemical enrichment. Saro et al. (2006) analysed the galaxy population from simulations of galaxy clusters and concluded that a Salpeter IMF produces a colour-magnitude relation that, with the exception of the BCGs, is in reasonable agreement with observations. On the contrary, the stronger enrichment provided by a top-heavier IMF turns into too red galaxy colours.

In the following sections, the above results on the cosmological modelling of the ICM will be reviewed in more details, also critically discussing the possible presence of observational biases, which may alter the determination of relative abundances and, therefore, the inference of the IMF shape.

Different variants of the chemical evolution model have been implemented by different authors in their simulation codes. For instance, the above described model of chemical evolution has been implemented with minimal variants by Kawata and Gibson (2003a), Kobayashi (2004), who also included the effect of hypernova explosions, and Tornatore et al. (2007a), while Tornatore et al. (2007b) also included the effect of metal enrichment from low-metallicity (Pop III) stars. Raiteri et al. (1996) and Valdarnini (2003) also used a similar model, but neglected the contribution from low- and intermediate-mass stars. Mosconi et al. (2001) and Scannapieco et al. (2005) neglected delay times for SN II, assumed a fixed delay time for SN Ia and neglected the contribution to enrichment from low- and intermediate-mass stars.

Clearly, a delicate point in hydrodynamical simulations is deciding how metals are distributed to the gas surrounding the star particles. The physical mechanisms actually responsible for enriching the inter-stellar medium (ISM; e.g., stellar winds, blast waves from SN explosions, etc.) take place on scales which are generally well below the resolution of current cosmological simulations of galaxy clusters (see Schindler and Diaferio 2008—Chap. 17,

this issue). For this reason, the usually adopted procedure is that of distributing metals according to the same kernel which is used for the computation of the hydrodynamical forces, a choice which is anyway quite arbitrary. Mosconi et al. (2001) and Tornatore et al. (2007a) have tested the effect of changing in different ways the weighting scheme to distribute metals and found that final results on the metal distribution are generally rather stable. Although this result is somewhat reassuring, it is clear that this warning on the details of metal distribution should always be kept in mind, at least until our simulations will have enough resolution and accurate description of the physical processes determining the ISM enrichment (Schindler and Diaferio 2008—Chap. 17, this issue).

3 Global Abundances

As a first step in our overview of the results of the ICM enrichment, we discuss a few general features of the enrichment pattern, which are related to the underlying model of chemical evolution. As a first example, we show in Fig. 3 maps of the fractional contribution of SN II to the global metallicity obtained from SPH simulations of a galaxy cluster having a temperature $\simeq 3$ keV (Tornatore et al. 2007a); the simulations include a self-consistent description of gas cooling, star formation and chemical evolution. The left and the right panels show the results based on assuming a Salpeter and a top-heavy AY IMFs, respectively. The clumpy aspect of these maps reveal that the products of SN II and SN Ia are spatially segregated. Since SN II products are released over a relatively short time-scale, they are preferentially located around star forming regions. On the other hand, SN Ia products are released over a longer time scale, which is determined by the life-times of the corresponding progenitors. As a consequence, a star particle has time to move away from the region, where it was formed, before SN Ia contribute to the enrichment. In this sense, the different spatial pattern of SN Ia and SN II products also reflects the contribution that a diffuse population of intra-cluster stars (e.g., Arnaboldi 2004 for a review) can provide to the ICM enrichment. Also, using a top-heavier IMF turns into a larger number of SN II (see also Fig. 2), with a resulting increase of their relative contribution to the metal production.

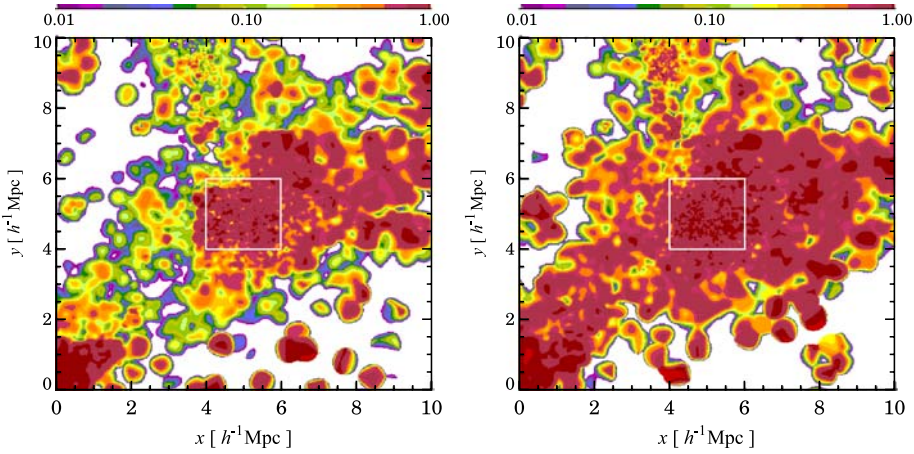


Fig. 3 Maps of the fractional contribution of SN II to the global metallicity (from Tornatore et al. 2007a). The *left panel* refers to the Salpeter IMF, while the *right panel* is for the AY top-heavy IMF. The white square is centred on the minimum of the cluster gravitational potential and its size is about twice the value of R_{200}

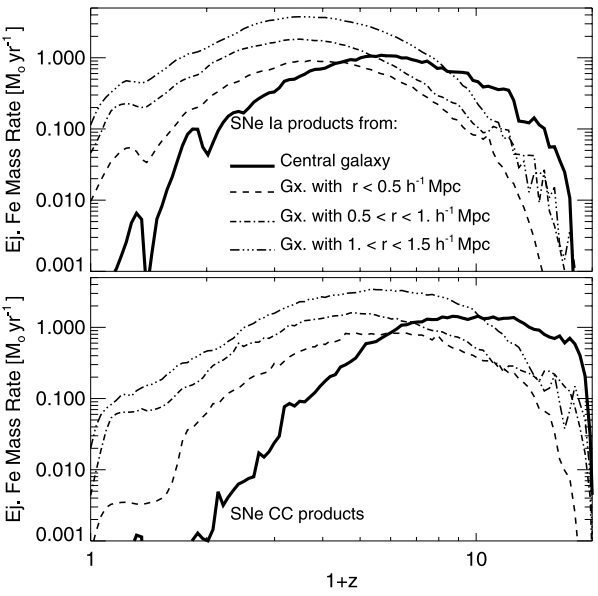


Fig. 4 Iron mass ejection rate as a function of redshift for SN Ia (*top panel*) and SN II (*bottom panel*), as predicted by the semi-analytical model by Cora (2006). The thick solid line is for the products released by the BCG, while the other curves are for galaxies lying within different cluster-centric distances, as computed at $z = 0$, after excluding the BCG contribution

To further illustrate the different timing of the enrichment from SN Ia and SN II, we show in Fig. 4 the rate of the iron mass ejection from the two SN populations, as predicted by a SAM model of galaxy formation, coupled to a non-radiative hydrodynamical simulation (Cora 2006). As expected, the contribution from SN II (lower panel) peaks at higher redshift than that of the SN Ia, with a quicker decline at low z, and closely traces the star formation rate (SFR). As for the SN Ia, their rate of enrichment is generally given by the convolution of the SFR with the lifetime function. As such, it is generally too crude an approximation to compute the enrichment rate by assuming a fixed delay time with respect to the SFR. Also interesting to note from this figure is that everything takes places earlier in the BCG (solid curves), than in galaxies at larger cluster-centric distances.

A crucial issue when performing any such comparison between observations and simulations concerns the definition of metallicity. From observational data, the metallicity is computed through a spectral fitting procedure, by measuring the equivalent width of an emission line associated to a transition between two heavily ionised states of a given element. In this way, one expects that the central cluster regions, which are characterised by a stronger emissivity, provides a dominant contribution to the global spectrum and, therefore, to the observed ICM metallicity. The simplest proxy to this spectroscopic measure of the ICM metallicity is, therefore, to use the emission-weighted definition of metallicity:

$$Z_{\rm ew} = \frac{\sum_i Z_i m_i \rho_{g,i} \Lambda(T_i, Z_i)}{\sum_i m_i \rho_{g,i} \Lambda(T_i)}. \tag{10}$$

In the above equation, Z_i, m_i, $\rho_{g,i}$ and T_i are the metallicity, mass, density and temperature of the i-gas element, with the sum being performed over all the gas elements lying within the cluster extraction region. Furthermore, $\Lambda(T, Z)$ is the metallicity dependent cooling function, which is defined so that $n_{\rm e} n_{\rm H} \Lambda(T, Z)$ is the radiated energy per unit time and unit volume from a gas element having electron number density $n_{\rm e}$ and hydrogen number density $n_{\rm H}$ (e.g., Sutherland and Dopita 1993). Since both simulated and observed metallicity radial profiles are characterised by significant negative gradients, we expect the "true"

mass-weighted metallicity,

$$Z_{\rm mw} = \frac{\sum_i Z_i m_i}{\sum_i m_i}, \tag{11}$$

to be lower than the "observed" emission-weighted estimate. Indeed, mock spectral observations of simulated clusters have shown that the above emission-weighted definition of metallicity is generally quite close to the spectroscopic value (Kapferer et al. 2007a; Rasia et al. 2007), at least for iron. On the other hand, Rasia et al. have also shown that abundances of other elements may be significantly biased. This is especially true for those elements, like oxygen, whose abundance is measured from transitions taking place at energies at which it is difficult to precisely estimate the level of the continuum in a hot system, due to the limited spectral resolution offered by the CCD on-board Chandra and XMM-Newton.

Based on a semi-analytical model of galaxy formation, Nagashima et al. (2005) carried out a comparison between the global observed and predicted ICM enrichment for different elements. In their analysis, these authors considered both the case of a "standard" Salpeter IMF and the case in which a top-heavier IMF is used in starbursts triggered by galaxy mergers. The comparison with observations, as reported in Fig. 5, led Nagashima et al. (2005) to conclude that the top-heavier IMF is required to reproduce the level of enrichment traced by different elements. All the abundances are predicted to be almost independent of temperature, at least for $T_X \gtrsim 1$ keV (this temperature corresponds to the limiting halo circular velocity, above which gas and metals ejected by superwinds are recaptured by the halo). While this result agrees with the observational points shown in the figure, they are at variance with respect to the significant temperature dependence of the iron abundance found by Baumgartner et al. (2005).

A different conclusion on the shape of the IMF is instead reached by other authors. For instance, Fabjan et al. (2008) analysed a set of hydrodynamical simulations of galaxy clusters, which include a chemical evolution model. In Fig. 6 we show the result of this analysis for the comparison between the simulated temperature dependence of iron (left panel) and sulphur (right panel) and the results of XMM-Newton observations by de Plaa et al. (2007) (see also Werner et al. 2008—Chap. 16, this issue). As for the iron abundance, there is a tendency for simulations with a Salpeter IMF to overproduce this element, by about 30 per cent. Using a AY top-heavy IMF turns into a much worse disagreement, with an excess of iron in simulations by a factor 2–3. It is clear that this iron overabundance could be mitigated in case simulations include some mechanism for diffusion and/or mixing of metals, which decreases the abundance at small cluster-centric radii. Of course, this decrease should be such to reproduce the observed metallicity profiles (see below). As for the [S/Fe] relative abundance,[1] the difference between the two IMFs is much smaller, since the two elements are produced in similar proportions by different stellar populations, and, within the relatively large observational uncertainties, they agree with data.

The comparison between the results of Figs. 5 and 6 is only one example of the different conclusions that different authors reach on the IMF shape from the study of the enrichment of the ICM. Tracing back the origin of these differences is not an easy task. Besides the different descriptions of the relevant physical processes treated in the SAMs and in the chemo-dynamical simulations, a proper comparison would also require using exactly the

[1] Here and in the following, we follow the standard bracket notation for the relative abundance of elements A and B: $[A/B] = \log(Z_A/Z_B) - \log(Z_{A,\odot}/Z_{B,\odot})$.

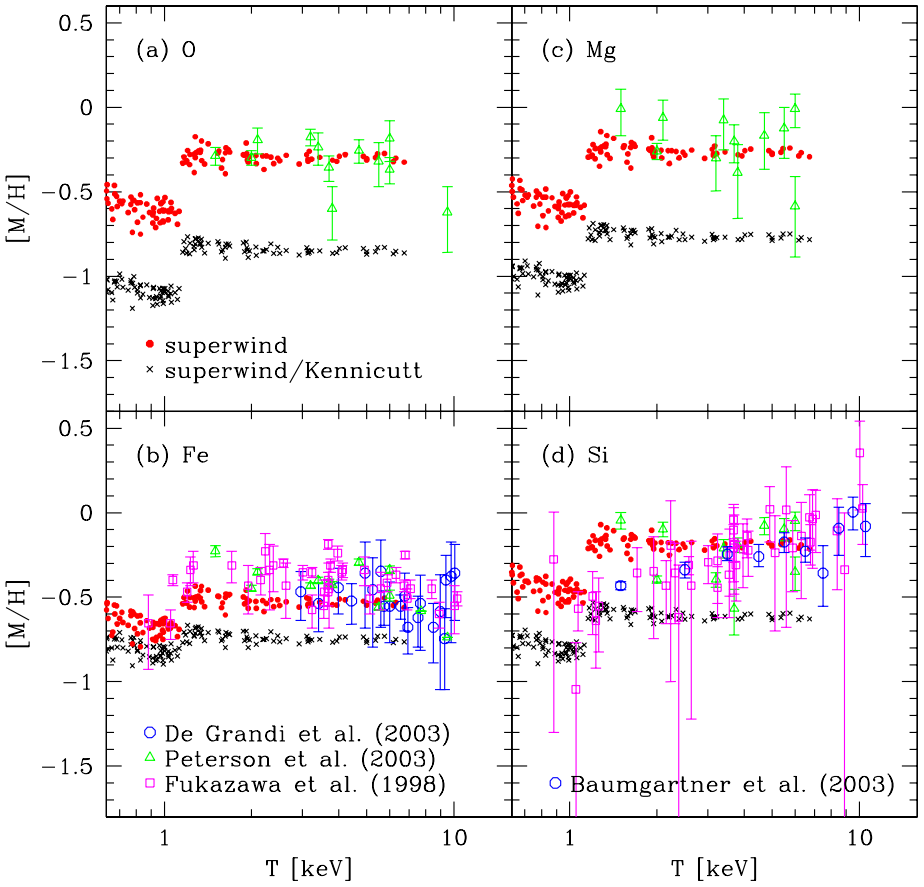


Fig. 5 The metal abundances of the ICM, as predicted by a semi-analytical model of galaxy formation (from Nagashima et al. 2005). Each panel shows a different element: (**a**) [O/H], (**b**) [Fe/H], (**c**) [Mg/H] and (**d**) [Si/H]. Model predictions are shown by *dots* for a top-heavy IMF, and by *crosses* for the Kennicutt IMF. Symbols with error bars correspond to observational data, which have been taken from De Grandi et al. (2004), Peterson et al. (2003), Fukazawa et al. (1998) and Baumgartner et al. (2005) (see Nagashima et al. 2005 for details). All abundances have been rescaled to the Solar abundances by Grevesse and Sauval (1998)

same model of chemical evolution and the same sets of yields for different stellar populations. We do recommend that all papers aimed at modelling the ICM enrichment should carefully include a complete description of the adopted chemical evolution model.

4 Metallicity Profiles

Observations based on ASCA (e.g. Finoguenov et al. 2001) and Beppo-SAX (De Grandi et al. 2004) satellites have revealed for the first time the presence of significant gradients in the profiles of ICM metal abundances. The much improved sensitivity of the XMM-Newton and Chandra satellites are now providing much more detailed information on the spatial distributions of different metals, thus opening the possibility of shedding light on the

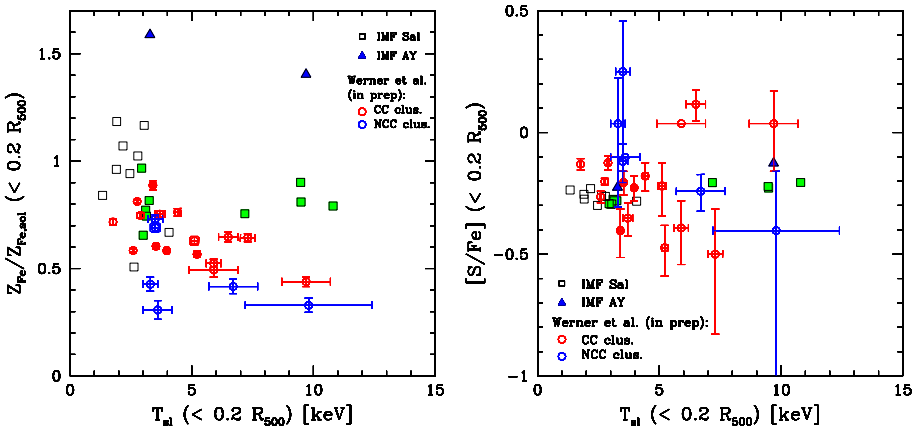


Fig. 6 *Left panel*: the iron abundance within $0.2r_{500}$ as a function of the cluster temperature (here r_{500} is defined as the radius encompassing an overdensity of $500\rho_c$, with ρ_c the cosmic critical density). The *open squares* are for simulated clusters analysed by Fabjan et al. (2008), which assume a Salpeter IMF, while the *filled triangles* are for two simulated clusters which assume the top-heavy AY IMF. *Open and filled circles* with error bars are for cool-core and non cool-core clusters, respectively, observed with XMM-Newton by de Plaa et al. (2007) (see also Werner et al. 2008—Chap. 16, this issue). *Right panel*: the same as in the *left panel*, but for the abundance of sulphur relative to iron. All abundances have been rescaled to the Solar abundances by Grevesse and Sauval (1998)

past history of star formation in cluster galaxies and on the gas-dynamical processes which determine the metal distribution.

A detailed study of the metal distribution in the ICM clearly requires resorting to the detailed chemo-dynamical approach offered by hydrodynamical simulations. In principle, a number of processes, acting at different scales, are expected to play a role in determining the spatial distribution of metals: galactic winds (e.g. Tornatore et al. 2004; Romeo et al. 2006), transport by buoyant bubbles (Roediger et al. 2007; Sijacki et al. 2007), ram-pressure stripping (e.g., Kapferer et al. 2007b), diffusion by stochastic gas motions (Rebusco et al. 2005), sinking of highly enriched low-entropy gas (Cora 2006), etc. It is clear that modelling properly all such processes represents a non-trivial task for numerical simulations, which are required both to cover a wide dynamical range and to provide a reliable description of a number of complex astrophysical processes (Schindler and Diaferio 2008—Chap. 17, this issue).

As an example, in the left panel of Fig. 7 we show the result of the analysis by Kapferer et al. (2007b), which was aimed at studying the relative role played by galactic winds and ram-pressure stripping in determining the metallicity profiles. In these simulations, based on an Eulerian grid code, these authors use a semi-analytical model to follow the formation and evolution of galaxies, and described the evolution of a global metallicity, instead of following different metal species. It is interesting to see how different mechanisms contribute in the simulations to establish the shape of the metallicity profile (see also Schindler and Diaferio 2008—Chap. 17, this issue). In particular, ram-pressure stripping turns out to be more important in the central regions, where the pressure of the ICM is higher. On the other hand, galactic winds are more effective at larger radii, where they can travel for larger distances, thanks to the lower ICM pressure.

As a further example, we show in the right panel of Fig. 7 the effect of AGN feedback on the distribution of metals in both gas and stars from the cluster simulations by Sijacki et al.

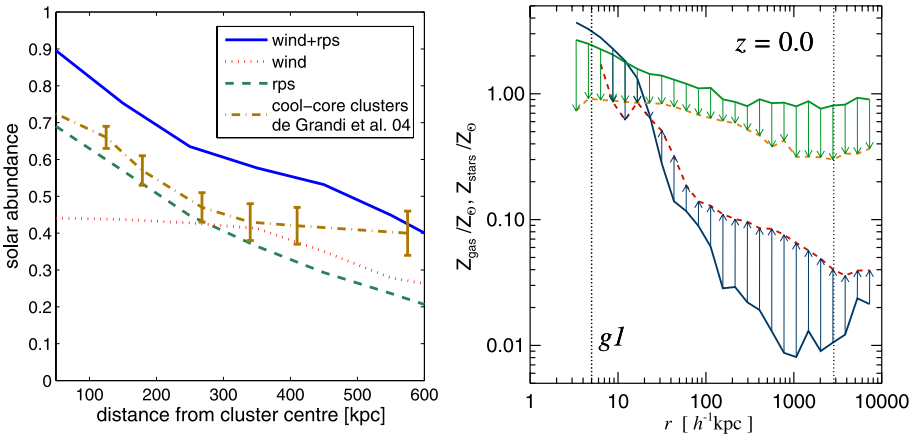


Fig. 7 *Left panel*: contribution from ram-pressure stripping and from galactic winds to the resulting iron metallicity profile of a simulated cluster (from Kapferer et al. 2007b). Data points are from the Beppo-SAX observations analysed by De Grandi et al. (2004). *Right panel*: the effect of AGN feedback on the profiles of mass-weighted gas metallicity and stellar metallicity with simulations of a galaxy cluster (from Sijacki et al. 2007). *Continuous lines* show the metallicity profiles without AGN heating, while *dashed lines* are for the runs with AGN feedback. The *arrows* illustrate how the ICM metallicity is affected by the AGN feedback. The *vertical dotted lines* mark the resolution limit of the simulations and the virial radius of the simulated cluster

(2007). These authors included in their SPH simulations a model for black hole growth and for the resulting energy feedback. Since these simulations do not include a detailed model of chemical evolution, the resulting metallicity cannot be directly compared to observations. Still, this figure nicely illustrates how the effect of AGN feedback is that of changing the share of metals between the stars and the gas. Due to the quenching of low-redshift star formation, a smaller amount of metals are locked back in stars, with a subsequent reduction of their metallicity. Furthermore, the more efficient metal transport associated to the AGN feedback causes a significant increase of the ICM metallicity in the outer cluster regions, thereby making the abundance profile shallower.

Since iron is the first element for which reliable profiles have been measured from X-ray observations, most of the comparisons performed so far with simulations have concentrated on its spatial distribution. As an example, we show in the left panel of Fig. 8 the comparison between the iron profiles measured by De Grandi et al. (2004) and results from the chemo-dynamical SPH simulations by Romeo et al. (2006). These authors performed simulations for two clusters having sizes comparable to those of the Virgo and of the Coma cluster, respectively. After exploring the effect of varying the feedback efficiency, they concluded that the resulting metallicity profiles are always steeper than the observed ones. A similar conclusion was also reached by Valdarnini (2003), based on SPH simulations of a larger ensemble of clusters, which used a rather inefficient form of thermal stellar feedback. A closer agreement between observed and simulated profiles of the iron abundance is found by Fabjan et al. (2008), see also Tornatore et al. (2007a), based on GADGET-2 SPH simulations, which include the effect of galactic winds, powered by SN explosions, according to the scheme presented by Springel and Hernquist (2003). The results of this comparison are shown in the right panel of Fig. 8. The solid black lines show the results for simulated clusters of mass in the range $\simeq (1-2) \times 10^{15} h^{-1} M_\odot$, based on a Salpeter IMF, while the dashed curve is for one cluster simulated with a top-heavy AY IMF. The observational data points are for a

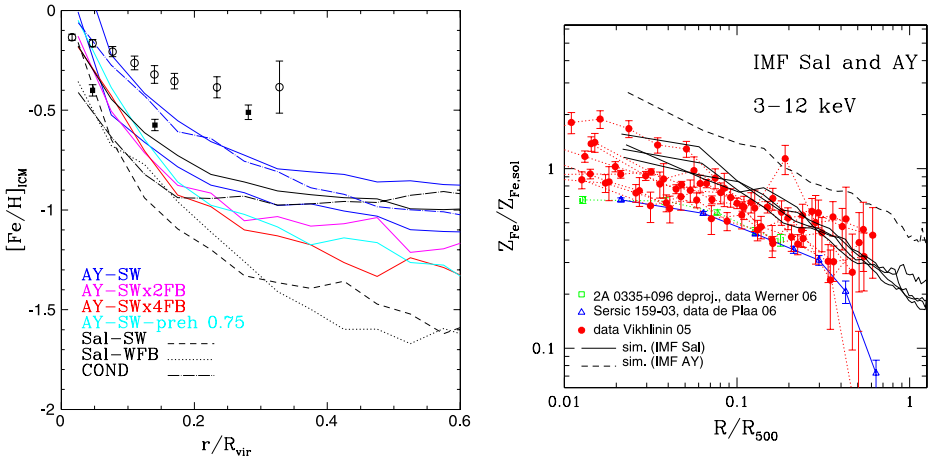


Fig. 8 *Left panel*: profiles of iron metallicity at $z = 0$ for different runs of two simulated clusters, obtained by changing the IMF (AY: Arimoto and Yoshii 1987; Salpeter 1955), the efficiency of the energy feedback and by including the effect of thermal conduction (COND; from Romeo et al. 2005). Observational data points are taken from De Grandi et al. (2004). *Right panel*: Comparison between observed and simulated profiles of iron abundance (from Fabjan et al. 2008). The continuous black curves are for four simulated clusters, having virial masses in the range $(1-2) \times 10^{15} h^{-1} \mathrm{M}_{\odot}$, assuming a Salpeter IMF. The black dashed line is for one of these clusters, simulated by assuming the AY top-heavy IMF. *Filled circles* with error bars show observational results from Chandra (Vikhlinin et al. 2005) for clusters with temperature above 3 keV, while the *open squares* and *triangles* are for the profiles of two $\simeq 3$ keV clusters observed with XMM-Newton (de Plaa et al. 2006; Werner et al. 2006). The iron abundance has been rescaled to the Solar abundances by Grevesse and Sauval (1998)

set of nearby relaxed clusters observed with Chandra and for two clusters with $T \simeq 3$ keV observed with XMM-Newton. This comparison shows a reasonable agreement between simulated and observed profiles, with the XMM-Newton profiles staying on the low side of the Chandra profiles, also with a shallower slope in the central regions. While this conclusion holds when using a Salpeter IMF, a top-heavy IMF is confirmed to provide too high a level of enrichment.

The different degree of success of the simulations shown in the two panels of Fig. 8 in reproducing observations can have different origins. The most important reason is probably due to the different implementations of feedback which, as shown in Fig. 7, plays a decisive role in transporting metals away from star-forming regions. Another possible source of difference may lie in the prescription with which metals are distributed around star particles. Once again, the details of the numerical implementations of different effects need to be clearly specified in order to make it possible to perform a meaningful comparison among different simulations. Finally, it is also interesting to note that clusters observed with Chandra seem to have steeper central profiles than those observed with XMM-Newton and Beppo-SAX. There is no doubt that increasing the number of clusters with reliably determined metallicity profiles will help sorting out the origin of such differences.

5 Evolution of the ICM Metallicity

After the pioneering paper based on ASCA data by Mushotzky and Loewenstein (1997), the increased sensitivity of the XMM-Newton and Chandra observatories recently opened the

possibility of tracing the evolution of the global iron content of the ICM out to the largest redshifts, $z \simeq 1.4$, where clusters have been secured (e.g., Balestra et al. 2007; Maughan et al. 2008). These observations have shown that the ICM iron metallicity decreases by about 50 per cent from the nearby Universe to $z \simeq 1$. This increase of metallicity at relatively recent cosmic epochs can be interpreted as being due either to a fresh production of metals or to a suitable redistribution of earlier-produced metals. Since SN Ia, which provide a large contribution to the total iron budget, can release metals over fairly long time scales, the increase of the iron abundance at $z < 1$ is not necessarily in contradiction with the lack of recent star formation in the bulk of cluster galaxies. On the one hand, recent analyses (Ettori 2005; Loewenstein 2006) have shown that the observed metallicity evolution is consistent with the observed SN Ia rate in clusters (Gal-Yam et al. 2002). On the other hand, a number of processes, such as stripping of metal enriched gas from merging galaxies (e.g., Calura et al. 2007) or sinking of highly enriched low-entropy gas towards the cluster central regions (Cora 2006) may also significantly contribute to the ICM metallicity evolution, through a redistribution of enriched gas. Indeed, the amount of accreted gas since $z = 1$ is at least comparable to that present within the cluster at that redshift. Therefore, the evolution of the ICM metallicity is expected to be quite sensitive to the enrichment level of the recently accreted gas. On the other hand, as noted by Renzini (1997), if ram-pressure stripping plays a dominant role in determining the ICM enrichment, one should then expect to observe more massive clusters, whose ICM has a higher pressure, to be more metal rich than poorer systems. In fact, observations do not show this trend and, if any, suggest a decrease of the iron abundance with the ICM temperature (e.g., Baumgartner et al. 2005).

Cora (2006) used a hybrid approach based on coupling a SAM model of galaxy formation to a non-radiative hydrodynamical simulation to study the gas dynamical effects which lead to the re-distribution of metals from the star forming regions. The results of her analysis are summarised in the metallicity maps shown in Fig. 9. Metals are produced inside small halos which at high redshift define the proto-cluster region. Since the gas in these halos has undergone no significant shocks by diffuse accretion, its entropy is generally quite low. This low-entropy highly enriched gas is therefore capable to sink towards the central regions of the forming cluster. As a consequence, this leads to an increase of the metallicity in the regions which mostly contribute to the emissivity, thereby causing a positive evolution of the observed metallicity.

Nagashima et al. (2005) used a SAM approach, not coupled to hydrodynamical simulations, to predict the evolution of the ICM metallicity. In this case, the evolution is essentially driven by the low-redshift star formation, since any redistribution of enriched gas is not followed by the model. The result of their analysis (left panel of Fig. 10) shows that a very mild evolution is predicted. A prediction of the evolution of the iron abundance from hydrodynamical simulations, which self-consistently follow gas cooling, star formation and chemical evolution, has been more recently performed by Fabjan et al. (2008). The results of their analysis are shown in the right panel of Fig. 10, where they are also compared to the observational results by Balestra et al. (2007). Also in this case, the results based on a Salpeter IMF are in reasonable agreement with observations, while a top-heavier IMF produces too high metallicities at all redshifts. Quite interestingly, at low redshift the run with top-heavy IMF shows an inversion of the evolution. This is due to the exceedingly high star formation triggered by heavily enriched gas, which locks back in stars a substantial fraction of the gas metal content in the cluster central regions. This highlights the crucial role played by star formation in both enriching the ICM and in locking back metals in the stellar phase. This further illustrates how the resulting metallicity evolution must be seen as the result of a delicate interplay of a number of processes, which all need to be properly described in a self-consistent chemo-dynamical model of the ICM.

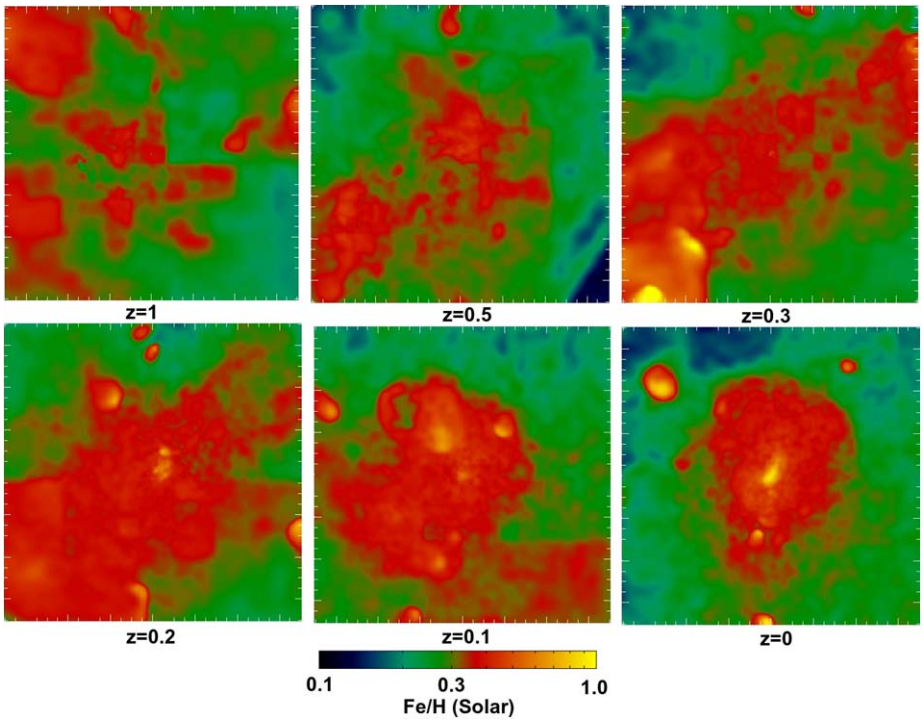


Fig. 9 Evolution of the iron abundance within 700 h^{-1} kpc from the centre of a simulated cluster at $z = 1, 0.5, 0.3, 0.2, 0.1, 0$, from *upper left to bottom right panel*. The maps show the projection of the mass-weighted iron abundance by number relative to hydrogen, Fe/H, with respect to the solar value (adapted from Cora 2006)

6 Properties of the Galaxy Population

The physical processes which determine the thermodynamical and chemical properties of the ICM are inextricably linked to those determining the pattern of star formation in cluster galaxies. For this reason, a successful description of galaxy clusters requires a multiwavelength approach, which is able at the same time to account for both the X-ray properties of the diffuse hot gas and of the optical/near-IR properties of the galaxy populations. In this sense, the reliability of a numerical model of the ICM chemical enrichment is strictly linked to its ability to reproduce the basic properties of the galaxy population, such as luminosity function and colour-magnitude diagram. While the study of the global properties of the galaxy populations has been conducted for several years with semi-analytical models of galaxy formation (De Lucia et al. 2004; Nagashima et al. 2005), it is only quite recently that similar studies have been performed with cosmological hydrodynamical simulations, thanks to the recent advances in supercomputing capabilities and code efficiency.

As we have already mentioned, hydrodynamical simulations treat the process of star formation through the conversion of cold and dense gas particles into collisionless star particles. Each star particle is then characterised by a formation epoch and a metallicity. Therefore, it can be treated as a SSP, for which luminosities in different bands can be computed by resorting to suitable stellar population synthesis models. Since the colours of a SSP are quite sensitive to metallicity, it is necessary for simulations to include a detailed chemical evolu-

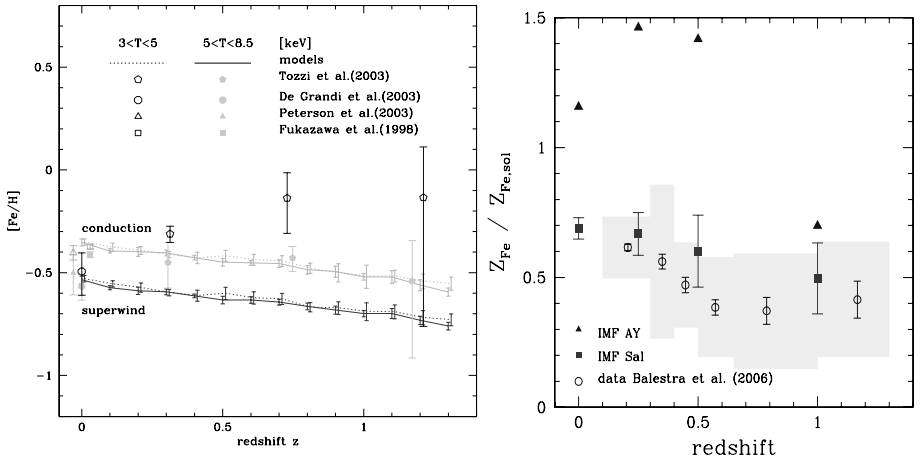


Fig. 10 The redshift evolution of the iron abundance. *Left panel*: predictions by different flavours of the SAM by Nagashima et al. (2005), based on a top-heavy IMF, compared with observational data from different sources, for clusters within two temperature ranges. Error bars on model predictions are the 1σ scatter over the ensemble of cluster formation histories. *Right panel*: comparison between observational data and full hydrodynamical simulations by Fabjan et al. (2008). *Open circles* come from the analysis of Chandra data by Balestra et al. (2007). The *shaded area* represents the r.m.s. scatter within each redshift interval, while the error bars are the 1σ confidence level. The *filled squares* are for a set of four simulated clusters, using a Salpeter IMF, with temperature $kT > 5$ keV, with error bars indicating the corresponding 1σ scatter. The *filled triangles* refer to only one cluster, out of four, simulated with the top-heavy IMF. In both panels, the values of the iron abundance have been rescaled to the Solar abundances by Grevesse and Sauval (1998)

tion model in order to reliably compare predictions on optical/near-IR properties of galaxies with observational data. This approach has been pursued by Kawata and Gibson (2003b), who performed simulations of elliptical galaxies and by Romeo et al. (2005) and Saro et al. (2006), who performed simulations of galaxy clusters. The results of these analyses are summarised in Fig. 11. In the left panels we show the position on the colour-magnitude diagram of the elliptical galaxies simulated by Kawata and Gibson (2003b). This plot clearly indicates that elliptical galaxies are generally much bluer than observed, as a result of an excess of recent star formation, which is not properly balanced by an efficient feedback mechanism. Indeed, it is only after excluding the contribution from recently formed stars that Kawata and Gibson (2003b) were able to obtain ellipticals with more realistic colours.

In the right panels we show the colour-magnitude relation (CMR) of the galaxy population of an extended set of clusters simulated by Saro et al. (2006) based both on a Salpeter IMF and on a top-heavy IMF. Quite apparently, the bulk of the galaxy population traces a CMR which is consistent with the observed one, at least for a Salpeter IMF. As for the top-heavy IMF, it predicts too red colours as a consequence of a too high metal content of galaxies. Also interesting to note, brighter redder galaxies are on average more metal rich, therefore demonstrating that the CMR corresponds in fact to a metallicity sequence (see also Romeo et al. 2005). However, a clear failure of these simulations is that the BCGs are always much bluer than observed. This result agrees with the exceedingly blue colour of simulated ellipticals shown in the left panels. This is due to an excess of recent star formation, which, in BCGs of $\sim 10^{15} M_\odot$ simulated clusters, can be as high as 500–1000 $M_\odot\ yr^{-1}$. As discussed by Borgani et al. (2008)—Chap. 13, this issue, the presence of overcooling in cluster simulations manifests itself with the presence of steep negative temperature gra-

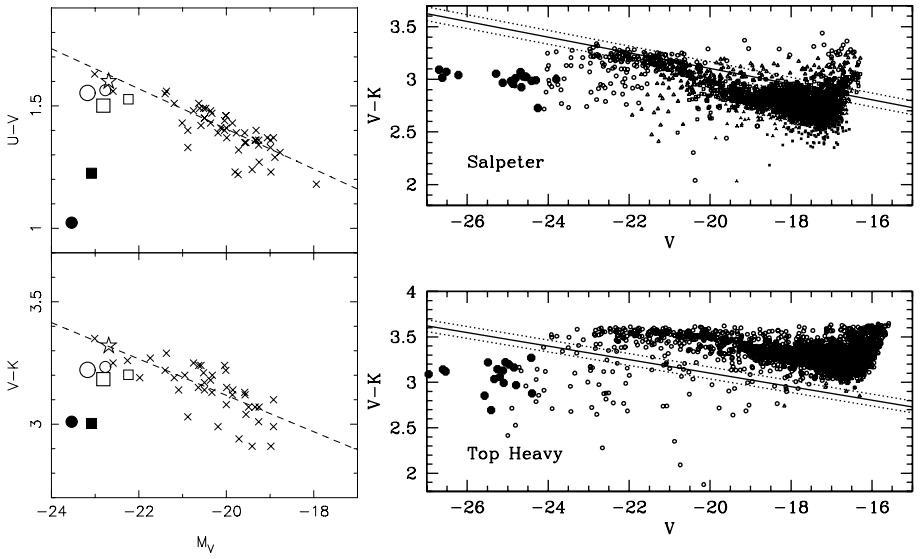


Fig. 11 *Left panel*: Colour-magnitude relations for the simulations of elliptical galaxies (Kawata and Gibson 2003b), based on a Salpeter IMF, compared to the observational results for the galaxies in Coma (Bower et al. 1992) (*crosses*) and for NGC 4472 (*open star*). The *filled square* and *circle* are the results for the reference runs, whereas the open symbols are when the contribution from stars younger than 2 or 8 Gyr are ignored. *Right panel*: the colour-magnitude relation for simulated clusters (*symbols*). *Upper and lower panels* refer to simulations performed with a Salpeter IMF and with the AY IMF, respectively (from Saro et al. 2006). The *big dots* are for the BCGs of the simulated clusters. The *straight lines* indicate the best-fitting relation and the scatter observed for the Coma galaxies

dients in the central regions. In this sense, wrong BCG colours and temperature profiles in core regions are two aspects of the same problem, i.e. the difficulty of balancing the cooling runaway within simulated clusters.

Since all these simulations include recipes for stellar feedback, although with different implementations and efficiencies, the solution to the overcooling in clusters should be found in some non-stellar feedback, such as that provided by AGN. Recent studies of the galaxy properties, based on semi-analytical models, have shown that an energy feedback of the sort of that expected from AGN, is able to suppress the bright end of the galaxy luminosity function, which is populated by the big ellipticals residing at the centre of clusters (e.g., Bower et al. 2006; Croton et al. 2006). There is no doubt that providing robust predictions on the properties of the galaxy population will be one of the most important challenges for simulations of the next generation to be able of convincingly describing the thermo- and chemo-dynamical properties of the hot gas.

7 Summary

The study of the enrichment pattern of the ICM and its evolution provides a unique means to trace the past history of star formation and the gas-dynamical processes, which determine the evolution of the cosmic baryons. In this overview, we discussed the concept of model of chemical evolution, which represents the pillar of any chemo-dynamical model, and presented the results obtained so far in the literature, based on different approaches. Restricting

the discussion to those methods which follow the chemical enrichment during the cosmological assembly of galaxies and clusters of galaxies, they can be summarised in the following categories.

(1) Semi-analytical models (SAMs) of galaxy formation (De Lucia et al. 2004; Nagashima et al. 2005): the production of metals is traced by the history of galaxy formation within the evolving DM halos and no explicit gas-dynamical description of the ICM is included. These approaches are computationally very cheap and provide a flexible tool to explore the space of parameters determining galaxy formation and chemical evolution. A limitation of these approaches is that, while they provide predictions on the global metal content of the ICM, the absence of any explicit gas dynamical description causes the lack of useful information on the spatial distribution of metals.
(2) SAMs coupled to hydrodynamical simulations. In this case, the galaxy formation is followed as in the previous approach, but the coupling with a hydrodynamical simulation allows one to trace the fate of the enriched gas and, therefore, to study the spatial distribution of metals. Different authors have applied different implementations of this hybrid approach, by focusing either on the role of the chemical evolution model (Cora 2006) or on the effect of specific gas-dynamical processes, such as ram-pressure stripping and galactic winds, in enriching the ICM (e.g., Kapferer et al. 2007b). The advantage of this approach is clearly that gas-dynamical processes are now included at some level. However, the galaxy formation process is not followed in a self-consistent way from the cooling of the gas during the cosmic evolution.
(3) Full hydrodynamical simulations, which self-consistently follow gas cooling, star formation and chemical evolution (Valdarnini 2003; Kawata and Gibson 2003b; Romeo et al. 2005; Tornatore et al. 2007a). Rather than being described by an external recipe, galaxy formation is now the result of the cooling and of the conversion of cold dense gas into stars, as implemented in the simulation code. The major advantage of this approach is that the enrichment process now comes as the result of the star formation predicted by the hydrodynamical simulation, without resorting to any external model. The bottleneck, in this case, is represented by the computational cost, which becomes prohibitively high if one wants to resolve the whole dynamic range covered by the processes of metal production and distribution, feedback energy release, etc. For this reason, these simulation codes also need to resort to sub-resolution models, which provide an effective description of a number of physical processes. However, there is no doubt that the ever improving supercomputing technology and code efficiency make, in perspective, this approach the way to study the past history of cosmic baryons.

As already emphasised, developing a reliable model of the cosmic history of metal production is a non-trivial task. This is due to the sensitive dependence of the model predictions on both the parameters defining the chemical evolution model (i.e., IMF, stellar life-times, stellar yields, etc.), and on the processes, such as ejecta from SN explosions and AGN, stripping, etc., which at the same time regulate star formation and transport metals outside galaxies. In the light of these difficulties, it is quite reassuring that the results of the most recent hydrodynamical simulations are in reasonable agreement with the most recent observational data from Chandra and XMM-Newton.

In spite of this success, cluster simulations of the present generation still suffer from important shortcomings. The most important of them is probably represented by the difficulty in regulating overcooling at the cluster centre, so as to reproduce at the same time the observed “cool cores” and the passively evolving massive ellipticals. Besides improving the description of the relevant astrophysical processes in simulation codes, another

important issue concerns understanding the possible observational biases which complicate any direct comparison between data and model predictions (Kapferer et al. 2007a; Rasia et al. 2007). In this respect, simulations provide a potentially ideal tool to understand these biases. Mock X-ray observations of simulated clusters, which include the effect of instrumental response, can be analysed exactly in the same way as real observational data. The resulting "observed" metallicity can then be compared with the true one to calibrate out possible systematics. There is no doubt that observations and simulations of the chemical enrichment of the ICM should go hand in hand in order to fully exploit the wealth of information provided by X-ray telescopes of the present and of the next generation.

Acknowledgements The authors thank ISSI (Bern) for support of the team "Non-virialized X-ray components in clusters of galaxies". We would like to thank Cristina Chiappini, Francesca Matteucci, Simone Recchi and Paolo Tozzi for a number of enlightening discussions. We also thank Florence Durret for a careful reading of the manuscript and Norbert Werner for providing the observational data points in Fig. 5 prior to publication. The authors acknowledge support by the PRIN2006 grant "Costituenti fondamentali dell'Universo" of the Italian Ministry of University and Scientific Research, by the Italian National Institute for Nuclear Physics through the PD51 grant, by the Austrian Science Foundation (FWF) through grants P18523-N16 and P19300-N16, and by the Tiroler Wissenschaftsfonds and through the UniInfrastrukturprogramm 2005/06 by the BMWF.

References

N. Arimoto, Y. Yoshii, Astron. Astrophys. **173**, 23 (1987)
M. Arnaboldi, in *Recycling Intergalactic and Interstellar Matter*. IAU Symp., vol. 217 (2004), p. 54
I. Balestra, P. Tozzi, S. Ettori et al., Astron. Astrophys. **462**, 429 (2007)
W.H. Baumgartner, M. Loewenstein, D.J. Horner, R.F. Mushotzky, Astrophys. J. **620**, 680 (2005)
G. Bertelli, A. Bressan, C. Chiosi, F. Fagotto, E. Nasi, Astron. Astrophys. Suppl. Ser. **106**, 275 (1994)
I.A. Bonnell, C.J. Clarke, M.R. Bate, Mon. Not. R. Astron. Soc. **368**, 1296 (2006)
S. Borgani, A. Diaferio, K. Dolag, S. Schindler, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9317-4
R.G. Bower, J.R. Lucey, R.S. Ellis, Mon. Not. R. Astron. Soc. **254**, 601 (1992)
R.G. Bower, A.J. Benson, R. Malbon et al., Mon. Not. R. Astron. Soc. **370**, 645 (2006)
F. Calura, F. Matteucci, P. Tozzi, Mon. Not. R. Astron. Soc. **378**, L11 (2007)
G. Chabrier, Publ. Astron. Soc. Pac. **115**, 763 (2003)
C. Chiappini, F. Matteucci, R. Gratton, Astrophys. J. **477**, 765 (1997)
A. Chieffi, M. Limongi, Astrophys. J. **608**, 405 (2004)
S.A. Cora, Mon. Not. R. Astron. Soc. **368**, 1540 (2006)
D.J. Croton, V. Springel, S.D.M. White et al., Mon. Not. R. Astron. Soc. **365**, 11 (2006)
S. De Grandi, S. Ettori, M. Longhetti, S. Molendi, Astron. Astrophys. **419**, 7 (2004)
G. De Lucia, G. Kauffmann, S.D.M. White, Mon. Not. R. Astron. Soc. **349**, 1101 (2004)
J. de Plaa, N. Werner, A.M. Bykov et al., Astron. Astrophys. **452**, 397 (2006)
J. de Plaa, N. Werner, J.A.M. Bleeker et al., Astron. Astrophys. **465**, 345 (2007)
A. Diaferio, G. Kauffmann, M.L. Balogh et al., Mon. Not. R. Astron. Soc. **323**, 999 (2001)
A. Diaferio, S. Schindler, K. Dolag, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9324-5
K. Dolag, S. Borgani, S. Schindler, A. Diaferio, A. Bykov, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9316-5
W. Domainko, M. Mair, W. Kapferer et al., Astron. Astrophys. **452**, 795 (2006)
S. Ettori, Mon. Not. R. Astron. Soc. **362**, 110 (2005)
S. Fabjan, L. Tornatore, S. Borgani, K. Dolag, F. Matteucci, Mon. Not. R. Astron. Soc. (2008 submitted)
A. Finoguenov, M. Arnaud, L.P. David, Astrophys. J. **555**, 191 (2001)
Y. Fukazawa, K. Makishima, T. Tamura et al., Publ. Astron. Soc. Jpn. **50**, 187 (1998)
A. Gal-Yam, D. Maoz, Mon. Not. R. Astron. Soc. **347**, 942 (2004)
A. Gal-Yam, D. Maoz, K. Sharon, Mon. Not. R. Astron. Soc. **332**, 37 (2002)
L. Greggio, Astron. Astrophys. **441**, 1055 (2005)
L. Greggio, A. Renzini, Astron. Astrophys. **118**, 217 (1983)
N. Grevesse, A.J. Sauval, Space Sci. Rev. **85**, 161 (1998)
K. Iwamoto, F. Brachwitz, K. Nomoto et al., Astrophys. J. Suppl. Ser. **125**, 439 (1999)

J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-y
W. Kapferer, T. Kronberger, J. Weratschnig, S. Schindler, Astron. Astrophys. **472**, 757 (2007a)
W. Kapferer, T. Kronberger, J. Weratschnig et al., Astron. Astrophys. **466**, 813 (2007b)
N. Katz, D.H. Weinberg, L. Hernquist, Astrophys. J. Suppl. Ser. **105**, 19 (1996)
D. Kawata, B.K. Gibson, Mon. Not. R. Astron. Soc. **340**, 908 (2003a)
D. Kawata, B.K. Gibson, Mon. Not. R. Astron. Soc. **346**, 135 (2003b)
C. Kobayashi, Mon. Not. R. Astron. Soc. **347**, 740 (2004)
P. Kroupa, Mon. Not. R. Astron. Soc. **322**, 231 (2001)
P. Kroupa, C.A. Tout, G. Gilmore, Mon. Not. R. Astron. Soc. **262**, 545 (1993)
R.B. Larson, Mon. Not. R. Astron. Soc. **301**, 569 (1998)
C. Lia, L. Portinari, G. Carraro, Mon. Not. R. Astron. Soc. **330**, 821 (2002)
M. Loewenstein, Astrophys. J. **648**, 230 (2006)
A. Maeder, G. Meynet, Astron. Astrophys. **210**, 155 (1989)
F. Mannucci, M. Della Valle, N. Panagia et al., Astron. Astrophys. **433**, 807 (2005)
P. Marigo, Astron. Astrophys. **370**, 194 (2001)
F. Matteucci, *The Chemical Evolution of the Galaxy*. Astrophysics and Space Science Library, vol. 253 (Kluwer, Dordrecht, 2003)
F. Matteucci, B.K. Gibson, Astron. Astrophys. **304**, 11 (1995)
F. Matteucci, S. Recchi, Astrophys. J. **558**, 351 (2001)
B.J. Maughan, C. Jones, W. Forman, L. Van Speybroeck, Astrophys. J. Suppl. Ser. **174**, 117 (2008)
C.F. McKee, E.C. Ostriker, Annu. Rev. Astron. Astrophys. **45**, 565 (2007)
M.B. Mosconi, P.B. Tissera, D.G. Lambas, S.A. Cora, Mon. Not. R. Astron. Soc. **325**, 34 (2001)
R.F. Mushotzky, in *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, ed. by J.S. Mulchaey, A. Dressler, A. Oemler (Cambridge University Press, Cambridge, 2004), p. 123
R.F. Mushotzky, M. Loewenstein, Astrophys. J. **481**, L63 (1997)
M. Nagashima, C.G. Lacey, C.M. Baugh, C.S. Frenk, S. Cole, Mon. Not. R. Astron. Soc. **358**, 1247 (2005)
K. Nomoto, K. Iwamoto, N. Nakasato et al., Nucl. Phys. A **621**, 467 (1997)
K. Nomoto, H. Umeda, I. Hachisu et al., in *Type Ia Supernovae, Theory and Cosmology*, ed. by J.C. Niemeyer, J.W. Truran (Cambridge University Press, Cambridge, 2000), p. 63
P. Padoan, Å. Nordlund, A.G. Kritsuk, M.L. Norman, P.S. Li, Astrophys. J. **661**, 972 (2007)
P. Padovani, F. Matteucci, Astrophys. J. **416**, 26 (1993)
J.R. Peterson, S.M. Kahn, F.B.S. Paerels et al., Astrophys. J. **590**, 207 (2003)
A. Pipino, F. Matteucci, S. Borgani, A. Biviano, New Astron. **7**, 227 (2002)
L. Portinari, C. Chiosi, A. Bressan, Astron. Astrophys. **334**, 505 (1998)
L. Portinari, A. Moretti, C. Chiosi, J. Sommer-Larsen, Astrophys. J. **604**, 579 (2004)
C.M. Raiteri, M. Villata, J.F. Navarro, Astron. Astrophys. **315**, 105 (1996)
E. Rasia, P. Mazzotta, H. Bourdin et al., Astrophys. J. (2007 in press). astro-ph/0707.2614
P. Rebusco, E. Churazov, H. Böhringer, W. Forman, Mon. Not. R. Astron. Soc. **359**, 1041 (2005)
A. Renzini, Astrophys. J. **488**, 35 (1997)
A. Renzini, in *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, ed. by J.S. Mulchaey, A. Dressler, A. Oemler (Cambridge University Press, Cambridge, 2004), p. 260
A. Renzini, M. Voli, Astron. Astrophys. **94**, 175 (1981)
E. Roediger, M. Brüggen, P. Rebusco, H. Böhringer, E. Churazov, Mon. Not. R. Astron. Soc. **375**, 15 (2007)
D. Romano, C. Chiappini, F. Matteucci, M. Tosi, Astron. Astrophys. **430**, 491 (2005)
A.D. Romeo, L. Portinari, J. Sommer-Larsen, Mon. Not. R. Astron. Soc. **361**, 983 (2005)
A.D. Romeo, J. Sommer-Larsen, L. Portinari, V. Antonuccio-Delogu, Mon. Not. R. Astron. Soc. **371**, 548 (2006)
P. Rosati, S. Borgani, C. Norman, Annu. Rev. Astron. Astrophys. **40**, 539 (2002)
E.E. Salpeter, Astrophys. J. **121**, 161 (1955)
A. Saro, S. Borgani, L. Tornatore et al., Mon. Not. R. Astron. Soc. **373**, 397 (2006)
C. Scannapieco, P.B. Tissera, S.D.M. White, V. Springel, Mon. Not. R. Astron. Soc. **364**, 552 (2005)
E. Scannapieco, L. Bildsten, Astrophys. J. **629**, L85 (2005)
S. Schindler, A. Diaferio, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9321-8
D. Sijacki, V. Springel, T. Di Matteo, L. Hernquist, Mon. Not. R. Astron. Soc. **380**, 877 (2007)
V. Springel, Mon. Not. R. Astron. Soc. **364**, 1105 (2005)
V. Springel, L. Hernquist, Mon. Not. R. Astron. Soc. **339**, 289 (2003)
M. Sullivan, D. Le Borgne, C.J. Pritchet et al., Astrophys. J. **648**, 868 (2006)
R.S. Sutherland, M.A. Dopita, Astrophys. J. Suppl. Ser. **88**, 253 (1993)
F.-K. Thielemann, D. Argast, F. Brachwitz et al., Nucl. Phys. A **718**, 139 (2003)
L. Tornatore, S. Borgani, F. Matteucci, S. Recchi, P. Tozzi, Mon. Not. R. Astron. Soc. **349**, L19 (2004)

L. Tornatore, S. Borgani, K. Dolag, F. Matteucci, Mon. Not. R. Astron. Soc. **382**, 1050 (2007a). astro-ph/0705.1921
L. Tornatore, A. Ferrara, R. Schneider, Mon. Not. R. Astron. Soc. **382**, 945 (2007b). astro-ph/0707.1433
A.V. Tutukov, L.R. Iungelson, in *Close Binary Stars: Observations and Interpretation*, ed. by M.J. Plavec, D.M. Popper, R.K. Ulrich IAU Symp., vol. 88 (1980), p. 15
R. Valdarnini, Mon. Not. R. Astron. Soc. **339**, 1117 (2003)
L.B. van den Hoek, M.A.T. Groenewegen, Astron. Astrophys. Suppl. Ser. **123**, 305 (1997)
A. Vikhlinin, M. Markevitch, S.S. Murray et al., Astrophys. J. **628**, 655 (2005)
G.M. Voit, Rev. Mod. Phys. **77**, 207 (2005)
N. Werner, J. de Plaa, J.S. Kaastra et al., Astron. Astrophys. **449**, 475 (2006)
N. Werner, F. Durret, T. Ohasi, S. Schindler, R.P.C. Wiersma, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9320-9
S.E. Woosley, T.A. Weaver, Astrophys. J. Suppl. Ser. **101**, 181 (1995)
L.R. Yungelson, M. Livio, Astrophys. J. **528**, 108 (2000)

Chapter 19
Future Instrumentation for the Study of the Warm-Hot Intergalactic Medium

Frits Paerels · Jelle Kaastra · Takaya Ohashi · Philipp Richter · Andrei Bykov · Jukka Nevalainen

Originally published in the journal Space Science Reviews, Volume 134, Nos 1–4.
DOI: 10.1007/s11214-008-9323-6

Abstract We briefly review capabilities and requirements for future instrumentation in UV- and X-ray astronomy that can contribute to advancing our understanding of the diffuse, highly ionised intergalactic medium.

Keywords Instrumentation: spectrographs · Instrumentation: photometers · Ultraviolet: general · X-rays: general · γ-rays: general

F. Paerels (✉)
Department of Astronomy and Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA
e-mail: frits@astro.columbia.edu

J. Kaastra
SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands

J. Kaastra
Sterrenkundig Instituut, Universiteit Utrecht, P.O. Box 80000, 3508 TA, Utrecht, The Netherlands

T. Ohashi
Department of Physics, School of Science, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, Japan

P. Richter
Institut für Physik, Universität Potsdam, Am Neuen Palais 10, 14469, Potsdam, Germany

A. Bykov
A.F. Ioffe Institute of Physics and Technology, Russian Academy of Sciences, 26 Polytekhnicheskaya, St. Petersburg 194021, Russia

J. Nevalainen
Observatory, University of Helsinki, P.O. Box 14, 00014 Tähtitornimäki, Finland

1 Introduction

A convergence of recent theoretical and observational work has generated rapidly growing interest in the physics of a highly ionised intergalactic medium (IGM). At low redshifts, this phase of the IGM likely holds the balance of the baryon mass density not accounted for by the mass densities in stars, diffuse gas in galaxies, the local Lyα forest, the Intracluster Medium (ICM) in galaxy clusters and groups, etc. (Fukugita and Peebles 2004). Model calculations suggest that in fact the baryon density in the IGM could indeed be comparable to the summed baryon densities of the known mass components, as would be required to reach the baryon density determined using independent arguments (big bang nucleosynthesis and the measured abundances of the light elements, and the measured fluctuation spectrum of the cosmic microwave background).

It is clear that the study of non-equilibrium phenomena in the low-redshift diffuse intergalactic medium, both within and outside bound structures, can produce a wealth of information on a large range of important astrophysical problems, ranging from the formation of galaxies to the generation of the first dynamically important magnetic fields. In this paper, we will attempt to summarise the requirements and prospects for significant progress on the observational study of these phenomena, considered in the light of planned or proposed instrumentation. We begin with a slightly abstract discussion of required and desired instrumental capabilities (spectral resolution, sensitivity, etc.), before we consider specific future or proposed instruments.

2 X-ray and UV Spectroscopy of the ICM and IGM

2.1 Imaging Spectroscopy (Emission Line Imaging)

Imaging X-ray spectroscopy will play a crucial role in the study of non-equilibrium phenomena in highly ionised diffuse media. High spectroscopic resolution is a requirement here, not only for the measurement (if possible) of velocity fields and the application of standard emission line plasma diagnostics, but also simply as a means to separate extragalactic emission from the bright and rich foreground emission from hot gas in and around our Milky Way Galaxy, the (time variable) contribution from geocoronal and heliospheric charge exchange excitation by the Solar wind (Wargelin et al. 2004) and to suppress the continuum background from unresolved point sources and non-X-ray instrumental backgrounds.

To start with the latter point (suppression of continuum backgrounds for the detection of line emission from highly ionised species in the WHIM), we need a rough estimate for the expected emission line intensity; the extragalactic background intensity in the soft X-ray band (below 653 eV, corresponding to O VIII Lyα at $z = 0$) can be inferred from an extrapolation of the smooth isotropic background above 1 keV. We assume that the emission is dominated by collisional excitation, ignoring the recombination-excited contribution, as well as the contribution from resonant excitation by the X-ray background continuum (Churazov et al. 2001). Even though the oxygen ionisation balance is probably largely determined by X-ray photoionisation, the line emission rate will respond strongly to heating of the gas, as it passes through the structure formation shocks. As far as oxygen is concerned, the WHIM is a hybrid medium, with the ionisation provided by photons, but the cooling (line emission) provided by electron collisions.

A simple estimate for the expected intensity in the resonance lines of O VII and O VIII is thus

$$\langle I \rangle \sim \frac{1}{4\pi} \langle n_e^2 \rangle l \, A \, f_i \, S_{ij}(T_e) \tag{1}$$

$$= \frac{1}{4\pi} C \langle n_e \rangle^2 l \, A \, f_i \, S_{ij}(T_e) \text{ photons cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}, \tag{2}$$

with A the absolute elemental abundance, f_i the fractions of atoms in the relevant ionisation state, l the path length, C the clumping factor ($C \equiv \langle n^2 \rangle / \langle n \rangle^2$). From explicit simulations (see for instance Davé et al. 2001), we estimate the clumping factor $C \sim 100$. The average should really be carried out explicitly over redshift as well, since the WHIM rapidly evolves at small redshift.

$S_{ij}(T_e) \equiv \langle \sigma(v) v \rangle$ is the thermal average of the collisional excitation cross section $\sigma(v)$ times the electron-ion velocity v. A useful discussion of thermal collisional excitation rates is given by Mewe (Mewe 1999; his Sect. 8.1; see also Kaastra et al. 2008—Chap. 9, this issue). *Very* roughly, for a H-like ion, the absolute collision cross section is of order $\pi a_0^2 / Z^2$ (a_0 is the Bohr radius, Z the nuclear charge). For electron temperatures small compared to the excitation energy (which is what we are dealing with here), a strong temperature dependence in $\langle \sigma(v) v \rangle$ results from the small, strongly temperature-dependent fraction of the electrons that have sufficient energy to excite the ions. The steep functional dependence of $S_{ij}(T_e)$ is almost entirely determined by the temperature dependence of this fraction. Mewe (1999) gives very useful numerical expressions for the excitation rates, as well as references to more detailed, precise data. We use these precise rates in our estimate for $\langle I \rangle$, since the analytical approximations, while nicely reproducing the functional shape of the temperature dependence, fail in absolute value by more than an order of magnitude for electron temperatures far below threshold.

For the O VIII Lyα and O VII resonance lines, a very rough estimate for the rate is plotted as a function of electron temperature in Fig. 1. We have assumed one-tenth Solar abundance, $f_i = 0.5$ for both ions (at the low densities characteristic of the average WHIM, the ionisation fractions have to be calculated from the photoionisation balance and depend mostly on density rather than on electron temperature; strictly speaking, f_i should have been taken inside the average for the density), and $l = \frac{1}{3} c/H_0$, with H_0 the Hubble constant, as suggested by the simulations (the WHIM does not rise until after redshifts $z \sim 1$ (Davé et al. 2001), and the integral will of course be weighted towards the higher densities, at the smaller redshifts).

This faint line emission has to be detected against at least the extragalactic continuum background (assuming it is largely unresolved, and that the background is dominated by the astrophysical background). If we adopt an average background continuum intensity of ≈ 40 photons cm^{-2} s^{-1} sr^{-1} keV^{-1} at $E \sim 0.5$ keV (see McCammon and Sanders 1990, their Fig. 12), we need an energy resolution of order $\Delta E \lesssim 2$–3 eV at 0.5 keV in order to keep the continuum background in a spectral resolution element comparable to the emission line count per resolution element. Obviously, higher contrast in emission line images requires higher spectral resolution.

The soft background ($E \lesssim 1$ keV) is largely dominated by the foreground emission from our Galaxy and its immediate environment. As was dramatically demonstrated by the Wisconsin/Goddard X-ray Quantum Calorimeter (XQC) rocket experiment (McCammon et al. 2002), most of this 'foreground' emission is in emission lines, dominated by the O VII and O VIII $n = 1$–2 resonance lines. But normalising simple models for the emission from $T_e \sim 10^6$ K, Solar abundance gas to the measured emission line intensities indicates that

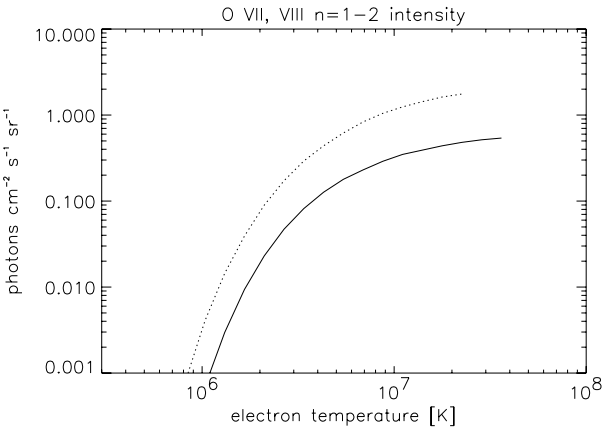


Fig. 1 Average photon count rates from the WHIM in the O VII $n = 1$–2 triplet (*dotted line*) and O VIII Lyα (*solid line*) lines, as a function of electron temperature. See text for explanation

there will be numerous weaker foreground lines, between which the weak redshifted, extragalactic emission lines will have to be detected. This will likely be impossible with data that does not at least have a resolution of $\Delta E \lesssim 2$ eV. For an impression of what the foreground spectrum might look like, see Fig. 13 in Kaastra et al. (2008)—Chap. 9, this issue. Note that due to the complexity of this foreground, convincing detection and interpretation of the extragalactic line emission will likely depend on a sophisticated simultaneous analysis of the Galactic emission.

As far as spectroscopic diagnostics are concerned, at $\Delta E \lesssim 4$–5 eV, the important He-like O VII $n = 1$–2 triplet is resolved. Resolving the thermal width of oxygen emission lines will require a very high resolving power $\mathcal{R} \gtrsim 1.3 \times 10^4 (T/10^6\ \mathrm{K})^{-1/2}$. More likely, the velocity dispersion observed in a finite spatial volume will be dominated by turbulence and bulk motion in the WHIM, with characteristic amplitude probably several 100 km s^{-1}, which is resolved at $\Delta E \lesssim 0.2$ eV in observations in which the individual WHIM filaments are spatially just resolved.

The predicted emission line intensities are very small, and an equally important instrument characteristic is the grasp (product of field of view Ω and effective area A) of the imaging spectrometer. With an intensity of ~ 0.1 photons cm^{-2} s^{-1} sr^{-1}, one gets less than 10 counts in 10^6 s of exposure if the grasp is smaller than $A\Omega \sim 10^{-4}$ cm^2 sr, or 0.3 cm^2 deg^2. For an effective mapping and true physical characterisation of the WHIM (as opposed to a single detection experiment), one probably needs at least two orders of magnitude larger grasp.

From the large scale structure/hydrodynamics simulations, the characteristic cross section of the WHIM filaments is about 10 Mpc, which spans an angle of $\sim 10^{-2}$ rad, or half a degree, at redshift $z = 0.3$. To purely resolve the WHIM structure, the angular resolution requirements therefore appear to be modest; instead, you need a large field of view of order at least a square degree to effectively map the structure. But low angular resolution, apart from contaminating the signal with unresolved point sources, will also introduce ambiguity in the precise location of the emitting ions with respect to other components of the large scale structure. At angular resolution $\Delta\theta \gtrsim 1$–2 arcmin, one can no longer distinguish truly diffuse intergalactic metal line emission, from emission due to small-scale galactic winds (and the interpretation of the data would involve assuming an explicit, complete model for the galactic feedback on the IGM). Finally, in addition to angular resolution requirements, any proposed optical system will also have to satisfy a contrast (or dynamic range) requirement: if too large a fraction of the photons are reflected or scattered far from the image core (i.e.

large wings on the point source response function), bright clusters in an image will outshine the faint WHIM emission. This may be a non-negligible problem, since the core densities of clusters and the average densities of the WHIM differ by a factor $\gtrsim$ 10–30, which result in a surface brightness contrast of a factor of up to 1000.

At the time of writing, a number of CCD-based imaging experiments, on the *Chandra*, *XMM-Newton*, and *Suzaku* observatories, are producing data. The spectral resolution is limited to $\Delta E \gtrsim 50$ eV, which makes it possible to discern and partially resolve C VI, N VII, O VII and O VIII, and Fe XVII line emission, but not without a model for contaminating fore- and backgrounds. Much more promising are imaging arrays of cryogenic spectrometers (microcalorimeters and Super Conducting Tunnelling Junctions, STJ), and such instruments are part of the baseline for the *Constellation-X* and *XEUS* observatories (for general descriptions of the operational principles of these types of detector, see for instance Enss 2005). In addition, dedicated soft X-ray, wide field microcalorimeter-based experiments have been proposed with the explicit aim of finding and characterising the WHIM (*DIOS, NEW, EDGE*). A brief description of these future instruments follows.

Both *Constellation-X*[1] (NASA), and *XEUS*[2] (ESA) are designed as general facilities, with a wide energy band, up to 10 keV (and a hard X-ray telescope above 10–15 keV, on *Constellation-X*), and, consequently, a rather long focal length and relatively small field of view (10 m focal length and F/8 focal ratio for *Constellation-X*; 35 m focal length for *XEUS* and $\sim$ F/10 focal ratio for *XEUS*). In the *Constellation-X* focal plane, an array of microcalorimeters provides imaging spectroscopy. In the baseline designs, the microcalorimeters are of the Transition Edge Sensor (TES) variety, with a design energy resolution goal of $\Delta E = 2$ eV. The field of view is approximately 5 arcmin in diameter, and the spectrometer has an effective area of ≈ 1 m^2 at sub-keV energies, achieved by flying four identical telescopes in parallel. *Constellation-X* also has a novel reflection grating spectrometer, which can be applied to high resolution spectroscopy of point sources. *XEUS* has been designed with two different types of cryogenic spectrometers in mind, Superconducting Tunnelling Junctions and TES microcalorimeters; the latter are currently the baseline. The energy resolution goals are 1 eV (STJ) and 2 eV (TES). The field of view of the cryogenic spectrometer will be approximately 0.75 arcmin diameter, with an effective area ~ 5 m^2 at low energies. Note that all cryogenic spectrometers need extensive blocking of UV/optical/infrared radiation, and the possible beam filters all have considerable low energy X-ray photoelectric absorption, resulting in sensitivities that steeply decline with decreasing photon energy. The effective areas listed above therefore are no more than easy reference numbers; for detailed sensitivity calculations, be sure to consult the relevant observatory technical information.

From the above characteristics, it is clear that the cryogenic imaging spectrometers on *Constellation-X* and *XEUS* will have sufficient energy resolution, but that their efficiency is limited. The relevant grasps are ~ 50 cm^2 deg^2, and ~ 6 cm^2 deg^2 for *Constellation-X* and *XEUS*, respectively. Note that these figures do not include the transmission of the optical blocking filters, which lowers the grasps by factors of a few at the X-ray energies of interest. In addition, the fields of view are not ideally matched to the problem. We will probably see line emission from the WHIM in (possibly) most deep observations, but only from a series of individual lines-of-sight.

By contrast, there has also been a number of proposals for smaller missions, that aim at detection of the WHIM line emission with a dedicated wide field of view, short focal length,

[1] http://constellation.gsfc.nasa.gov/.

[2] http://www.rssd.esa.int/index.php?project=XEUS&page=index.

high spectral resolution, soft X-ray imaging spectrometer. These characteristics (with the exception of the spectral resolution) are interlocking: a short focal length promotes sensitivity to faint emission against non-astrophysical background, allows a large field of view, but limits the energy response to low X-ray energies (which is ideal for study of the WHIM). Coupled with a high resolution imaging spectrometer, this is the natural design for study of extended diffuse faint soft X-ray line emission. We will therefore describe these mission concepts in somewhat greater detail.

The *Diffuse Intergalactic Oxygen Surveyor (DIOS)* concept originated in Japan (Ohashi et al. 2006). It is built around an imaging array of TES microcalorimeters, with an energy resolution of 2 eV, a 50 arcmin diameter field of view, an effective area in the O K band of > 100 cm^2, which gives a grasp of order 100 cm^2 deg^2. The mission lifetime should be > 5 yrs. With this grasp and field of view, the WHIM should clearly come into view; the total fraction of the baryonic mass density probed could be as high as 20 %.

A similar concept was studied by the Dutch: *NEW*, or the *Netherlands Explorer of the Web* (den Herder et al. 2006). In outline, it is very similar to *DIOS*. It has a 24^2 array of TES microcalorimeters, viewing a 1×1 deg^2 field of view at 2 eV energy resolution. The effective area is substantially larger than *DIOS*, with a 500 cm^2 baseline design in the O K band.

Other proposed missions incorporate the imaging spectroscopy of the high-ionisation metal emission lines with arrays of cryogenic spectrometers, but complement this with absorption line spectroscopy (*Estremo* and *EDGE*). For convenience, we will discuss these missions below, in the section on X-ray absorption spectroscopy.

Finally, the broad Hydrogen Lyα absorption lines discussed in Richter et al. (2008)—Chap. 3, this issue, trace the (small) neutral H content of the highly ionised medium. One can search for the corresponding faint Lyα emission lines, which would have to be done with a high-resolution UV imaging spectrometer. Work is proceeding on a balloon-borne experiment, to perform a first survey of line emission in the 2000–2200 Å band; this band is accessible from a balloon, and is limited by strong atmospheric absorption on either end (*FIREBALL*, or *F*aint *I*ntergalactic *R*edshifted *E*mission *BALL*oon[3]). With a proven spectrometer concept, a survey from space could then follow up at shorter wavelengths. The 2000–2200 Å window contains H Lyα λ1216 Å at $z = 0.7$, C IV λ1550 Å at $z = 0.36$, and O VI λ1033 Å at $z = 1.0$. The instrument, based on an echelle spectrograph, images approximately 200×200 arcsec^2, and can send the light of 20 individual, contiguous 7 arcsec wide strips through the spectrograph. The resolving power set by the 7 arcsec slit size is $\mathcal{R} = 8000$ in the first version of the instrument. The sensitivity is such that a diffuse Lyα intensity of ~ 500 photons cm^{-2} s^{-1} sr^{-1} can be detected in a single night, which corresponds to emission from regions of overdensity $\delta \sim 300$–1000.

2.2 Absorption Line Spectroscopy

2.2.1 Ultraviolet

As discussed by Richter et al. (2008)—Chap. 3, this issue, UV absorption spectroscopy of intervening hydrogen and metal absorbers currently represents the most sensitive method to study the statistical and physical properties of warm-hot intergalactic gas at low redshift. During the last years, UV spectra of several low-redshift QSOs obtained with the *Far Ultraviolet Spectroscopic Explorer* (*FUSE*) and the *Space Telescope Imaging Spectrograph*

[3] http://www.srl.caltech.edu/sal/igm_project.htm.

(*STIS*) aboard the *Hubble Space Telescope* (*HST*) have been used to study the properties of WHIM absorbers in great detail. By now, however, both instruments unfortunately are out of commission due to technical problems and therefore additional UV observations of the WHIM have to await future space-based UV instruments, which are discussed in the following.

While the spectral resolution of the former *FUSE* and *STIS* UV spectrographs ($\mathcal{R} \approx$ 20 000–45 000) were sufficient to reliably identify and analyse absorption features from the WHIM at low redshifts (in contrast to current X-ray spectrographs; see Richter et al. 2008—Chap. 3, this issue), the most important requirement for the next-generation UV spectrographs is a substantial gain in sensitivity. Only with more sensitive UV instruments will it be possible to achieve better statistics on the distribution and properties of WHIM absorbers along a large number of QSO sightlines at low and intermediate redshifts and to obtain spectra with higher signal-to-noise ratios. The latter aspect, in particular, is crucial to minimising systematic errors in the data analysis and interpretation of the UV absorption signatures.

Fortunately, the next UV spectrograph most likely will become available already in late 2008, when the *Cosmic Origins Spectrograph* (*COS*) will be installed on *HST* during the fifth and final service mission (SM−5) of *HST* (note that SM−5 is nothing but the reinstated SM−4 that was cancelled in 2004). SM−5 (flight STS−125) currently is scheduled for August 7, 2008 (see www.nasa.gov for regular updates on the NASA space shuttle launch schedule). *COS* is designed for high throughput, medium-resolution ($\mathcal{R} \sim 20\,000$) spectroscopy of point sources in the UV wavelength range. The *COS* instrument has two channels, a far-UV channel covering 1150–1775 Å, and a near-UV channel operating in the range 1750–3000 Å. The *COS* instrument has been built with maximum effective area as primary constraint. It will provide a gain in sensitivity of more than one order of magnitude compared to *STIS* and other previous spectrographs (e.g., *GHRS*) installed on the *HST*. *COS* is therefore expected to deliver a large number of high-quality spectra of low- and intermediate-redshift QSOs and AGN. These data will allow us to better understand the baryon budget and physical conditions in intervening O VI systems and broad Ly α absorbers (BLAs) and the space distribution of warm-hot intergalactic gas in relation to galactic structures. Note that it is also planned to repair *STIS* during the same service mission, so that there will be two powerful UV spectrographs with different spectral capabilities available at the *HST* at the same time.

Next to the *COS* and *STIS* instruments aboard the *HST*, there possibly will be another powerful UV instrument available at the beginning of the next decade. The Russian-led *World Space Observatory* (*WSO*) is an independent UV space telescope and is designed to represent an international follow-up of the *HST* mission. Next to Russia, other countries such as China, Spain, Italy, Germany are also involved in the *WSO* project and partly contribute with either money and/or instrumentation. The *WSO* will be equipped with a 1.7 m primary mirror, two high-resolution spectrographs with $\mathcal{R} \sim 50\,000$–$55\,000$ operating in the FUV and UV wavelength bands 1020–1720 Å, and 1740–3100 Å, and a low-resolution spectrograph ($\mathcal{R} \sim 2\,500$) for the range 1020–3100 Å (for details and further information see *WSO* homepage).[4] A big advantage of the *WSO* is that it represents an instrument entirely dedicated to the UV range. This will enable observers also to aim for long integration times of relatively faint (i.e. distant) UV background sources without facing competition with observers in other wavelength bands (as is the case for *COS* on *HST*, for instance). In addition,

[4] http://wso.inasan.ru.

WSO observations of intermediate-redshift QSOs will be particularly important to search for moderately redshifted WHIM EUV absorption from Ne VIII and Mg X (see Richter et al. 2008—Chap. 3, this issue, Sect. 2.2.1) that can be observed in the *WSO* FUV band. Consequently, the *WSO* will represent another very powerful UV instrument to study ultraviolet absorption signatures of the WHIM at low and intermediate redshift.

Next to *STIS*, *COS*, and the *WSO*, which by now represent approved science missions, there are various future mission concepts for space-based UV instruments, for instance as part of NASA's Origins Program.[5] Among these concepts, the *Baryonic Structure Probe* (PI: Kenneth Sembach, STScI) and the *High Orbit Ultraviolet-visible Satellite* (*HORUS*; PI: Jon Morse, Arizona State University) represent particularly interesting concepts for future UV observatories, as these instruments would be able to investigate warm-hot intergalactic gas in the ultraviolet in *both* absorption and emission.

2.2.2 X-ray

As discussed in Richter et al. (2008)—Chap. 3, this issue, detecting X-ray line absorption from the highly ionised light- and medium-Z elements in the WHIM (in practice, C, perhaps N, O, Ne, perhaps Fe) is challenging, due to the fact that the expected equivalent widths are very small, as well as to the fact that the absorption is expected to be sparse, at any spectrometer sensitivity currently foreseen. The absorption will be sparse in two senses: a single absorbing filament of WHIM gas will produce only one, or at most a few absorption lines; and on any line of sight to a bright, distant background continuum source we expect to be able to see only of order a few filaments at most. It is conceptually somewhat misleading to refer to the intergalactic X-ray absorption lines as the 'X-ray Forest' counterpart to the Lyα Forest (Hellsten et al. 1998). Nevertheless, X-ray absorption spectroscopy of the WHIM complements study of its line emission in very important ways: the sensitivity of absorption spectroscopy, for a given spectrometer sensitivity, only depends on the continuum flux of the background source, and, given sufficiently bright sources, studying the WHIM in absorption is potentially more sensitive to small column density filaments, as well as more distant filaments. Detecting the line *emission* from distant and/or low-density filaments will always be difficult. Finally, it is possible in principle to build (very) high resolution slitless grating spectrometers for point-source spectroscopy, and one could study the velocity structure (and hence the physical structure) of individual filaments, and the relation between, for instance, Lyα and O VI absorbers, and more highly ionised X-ray absorbers. Performing similarly high resolution diffuse emission line spectroscopy will probably be even much harder. On the other hand, given the sparseness of suitable backlighter sources, the absorption spectroscopy will most likely not reveal much about the full 3D topology of the WHIM, as the emission line imaging naturally does.

Detecting the weak X-ray absorption lines will require high spectral resolution and high sensitivity, the latter to ensure a sufficiently large number of photons per resolution element, as well as to make sure sufficiently weak (and therefore numerous) sources can be used, thereby covering both the sky as well as providing enough redshift depth. The appropriate scale for the spectral resolution is set by the thermal width of the lines. In the worst case (no large scale or turbulent velocity fields) the C, N, O, and Fe resonance lines will rapidly saturate, and the equivalent width will approximately stall at the value it attains at the thermal width. The thermal width, expressed as the Gaussian velocity dispersion, is $\sigma_v = \sqrt{kT/m_i} =$

[5] http://www.nasa.gov/home/hqnews/2004/jul/HQ-04246_mission_concepts.html.

$22(T/10^6\ \mathrm{K})^{1/2}$ km s^{-1} for oxygen ions. In Kaastra et al. (2008)–Chap. 9, this issue (their Fig. 9), we show the curve of growth for O VII and O VIII resonance lines; at the thermal width, these first saturate at equivalent width $W_\lambda \approx 4$ mÅ, which corresponds to a resolving power of 5000. The equivalent number for the resolution in energy measure is $\Delta E \approx 0.1$ eV.

In Richter et al. (2008)—Chap. 3, this issue, we reproduced a cumulative probability distribution, derived from large scale structure/hydrodynamics simulations, for the O VII and O VIII column densities. At equivalent width $W_\lambda = 4$ mÅ, the expectation value for the number of filaments per unit redshift producing that EW or larger on a random line of sight is approximately $dN/dz \approx 3$. A requirement for a future experiment could therefore be phrased as 'the capability to detect $\lesssim 4$ mÅ equivalent width absorption lines in multiple high redshift ($z \gtrsim 1$) continuum sources'.

The grating spectrometers on *Chandra* and *XMM-Newton* do not have sufficient resolution; they have $\Delta\lambda \approx 50$ mÅ at 20 Å (the *Chandra* HETGS has better resolution, but only very limited efficiency in the O K band). The results of searches on very bright targets with these instruments are described in Richter et al. (2008)—Chap. 3, this issue.

The imaging X-ray spectrometers discussed above can also be used for absorption spectroscopy, of course. The relatively coarse energy resolution ($\Delta E \sim$ 1–few eV) is not better than the resolving power we currently have with the grating spectrometers on *Chandra* and *XMM-Newton* ($\Delta\lambda = 0.05$ Å corresponds to $\Delta E = 1.5$ eV at 20 Å), and the improvement in sensitivity would come entirely from the larger effective area. In order to significantly improve our understanding of the WHIM by X-ray absorption spectroscopy with practicable cryogenic spectrometers, the effective area should therefore be appreciably larger than those of the *Chandra* and *XMM-Newton* grating spectrometers. Nevertheless, a survey based on moderate sensitivity spectroscopy of, roughly speaking, the extragalactic sources in the 4U catalog is not likely to produce much relevant information, beyond bare detection of the WHIM.

Even though the energy resolution of the microcalorimeter spectrometer envisioned for *XEUS* is not ideal for the intergalactic absorption spectroscopy, pointing 5 m^2 of spectroscopic effective area at a bright continuum source will of course produce a very high signal-to-noise continuum, in which absorption lines of equivalent width well below the formal spectroscopic resolution will be detectable in principle. The recent attempts with the grating spectrometers on *Chandra* and *XMM-Newton* under similar circumstances have illustrated how important a very detailed understanding of the continuum sensitivity of the spectrometer is; in fact, in practise this understanding could well be the ultimate limiting factor on the sensitivity, not the photon counting fluctuations or background characteristics. The goals for intergalactic absorption spectroscopy with *XEUS* have been summarised by Barcons (2003): a spectrum on each of about 50 AGN at redshifts $z \gtrsim$ few tenths, sensitive enough to detect oxygen resonance absorption lines of equivalent width ~ 0.4 eV.[6]

Using the microcalorimeter spectrometer on *Constellation-X* for absorption spectroscopy faces a similar problem, since the projected energy resolution is very similar to that for *XEUS*. But the baseline design for the observatory includes a reflection grating spectrometer with substantially higher spectral resolution (though smaller effective area; Cottam et al. 2006). The baseline design calls for a resolving power of at least 300 at 20 Å, with a goal of 3000. The effective area should exceed 1000 cm^2, so that the sensitivity should be at least a factor 10 higher than currently reachable with the Reflection Grating Spectrometer

[6]For a recent update, see the *XEUS* Science Requirements Document, 4th edition, http://www.rssd.esa.int/index.php?project=XEUS&page=Multimedia, p. 11.

on *XMM-Newton*. This is illustrated graphically on the *Constellation-X* website;[7] but note that the prominent absorption line in that figure (which presumably displays the capability of the *Constellation-X* grating spectrometer) is the O VII $n = 1$–2 resonance line at redshift $z = 0$ (from gas in a putative hot halo of our Galaxy and/or the intragroup medium of the Local group), which has an equivalent width of ≈ 13 mÅ (Williams et al. 2005; Rasmussen et al. 2007), a factor 3 larger than the strongest true intergalactic absorption lines we expect.

The problem of a sufficiently large population of suitably bright point sources at non-trivial redshifts can be solved by using a 'renewable' set of sources: gamma-ray burst afterglows. If one catches the afterglow early enough, the cumulative fluence over the course of the burst decay compares favourably with the fluence of the brightest extragalactic steady sources, observed over a similar time interval. There is no limit to the number of sources this way; moreover, GRB happen to originate on average at fairly high redshift ($z \gtrsim 1$), which is ideal for the WHIM problem. But in order to take advantage of these characteristics, one needs to be able to point at a GRB afterglow literally within a matter of minutes. The *Pharos*;[8] (Elvis and Fiore 2003) and *Estremo* (Piro et al. 2006) mission concepts build on the all-sky GRB monitoring and rapid repointing scheme first employed by *Swift*. *Pharos* employs a relatively short focal length, moderate angular resolution soft X-ray telescope to feed high reflectivity, high dispersion reflection gratings, of the so-called off-plane variety (see for instance McEntaffer et al. 2004), and reading out in the third spectral order. The spectrometer is designed to reach the thermal width of the oxygen absorption lines (~ 4 mÅ at 20 Å). *Estremo*, instead, relies on non-diffractive spectroscopy, with an array of TES microcalorimeters.

More recently, the *Estremo*, *NEW*, and *DIOS* ideas have tentatively been merged into a new proposal to ESA's *Cosmic Vision* program, under the name *EDGE* (Piro et al. 2007; *E*xplorer of *D*iffuse emission and *G*amma-ray burst *E*xplosions). The *EDGE* concept relies on rapid repointing to perform absorption spectroscopy on gamma-ray burst afterglows, with a TES-microcalorimeter-based spectrometer, with baseline energy resolution $\Delta E = 3$ eV. The same array of microcalorimeters views an area of 0.7×0.7 deg^2 of the sky, with an effective area of 1000 cm^2, which translates to a grasp of 500 cm^2 deg^2. With this instrument, the observatory will carry out an imaging soft X-ray emission line survey as part of its core mission; part of the available exposure time will be dedicated to imaging the WHIM on lines of sight for which there is a significant absorption spectrum. There is compelling scientific rationale for combining absorption and emission spectroscopy. If nothing else, determining both the optical depth and the emission measure in a given transition gives a rough measurement of linear size and ion density of the system; with assumptions on the atomic abundance, one can derive the total density. This, the largest and most ambitious of the various proposed WHIM experiments, will probe[9] a significant fraction of the phase plane of the low-density IGM at $z \lesssim 1$ (see Fig. 2).

2.3 Comparisons

We will attempt to summarise the capabilities of the specific X-ray instrumentation discussed above in a single graph. The sensitivity to detection of either weak emission or absorption lines will roughly scale as $(A\Omega/\Delta E)^{1/2}$, with A and Ω the effective area and field

[7] http://constellationx.nasa.gov/science/missing_baryons/whim.html.

[8] http://hea-www.harvard.edu/~elvis/Pharos_GSFC_Aug06.pdf.

[9] Or rather, could have probed; ESA did not select the concept for the Cosmic Vision program in the Fall of 2007.

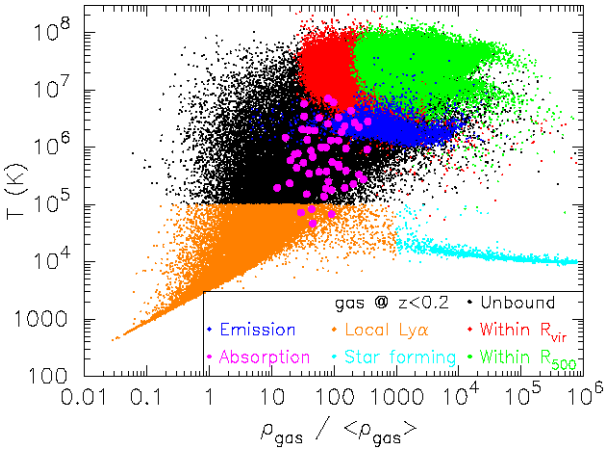


Fig. 2 Phase diagram of the intergalactic medium at $z < 0.2$. The dots are values of electron temperature and overdensity recorded in the computational volume of a large scale hydrodynamics simulation for $z < 0.2$; the various colours indicate different characteristic components of the diffuse gas: the warm, photoionised Lyα forest, the gas in virialised structures (groups and clusters), and the cool gas in (star-forming) galaxies. The WHIM-like points are *black* and *dark blue*. The *dark blue points* indicate the region of the density/temperature plane that will be probed in emission by the cryogenic imaging spectrometer on *EDGE*; the *purple dots* indicate the region that can be probed in absorption towards bright GRB afterglows in a 3 year core mission. From Piro et al. (2007)

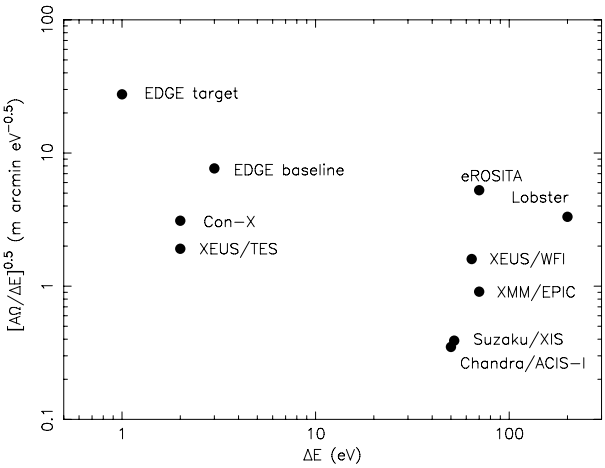


Fig. 3 A figure of merit for the detection of faint line absorption or emission from the WHIM, in the statistics-limited regime, the square root of the grasp divided $A\Omega$ by the spectral resolution ΔE, in units $(\mathrm{m}^2\,\mathrm{arcmin}^2\,\mathrm{eV}^{-1})^{0.5}$, versus spectral resolution ΔE, for a number of planned and currently operating X-ray imaging spectroscopic observatories

of view, and ΔE the spectral resolution, for the purely counting statistics-dominated case. For each of the instruments, we plot this figure of merit against the spectral resolution in Fig. 3. We argued that the requirements for detecting WHIM emission or absorption limit the viable experiments to the region $\Delta E < 2$ eV, $A\Omega > 30\,\mathrm{cm}^2\,\mathrm{deg}^2$, or $(A\Omega/\Delta E)^{1/2} > 2.3$ m arcmin $\mathrm{eV}^{-1/2}$. That shows that the future large observatories *XEUS* and *Constellation-X* are right at the lower end of the requirements; a large dedicated experiment such as *EDGE* has the required capability.

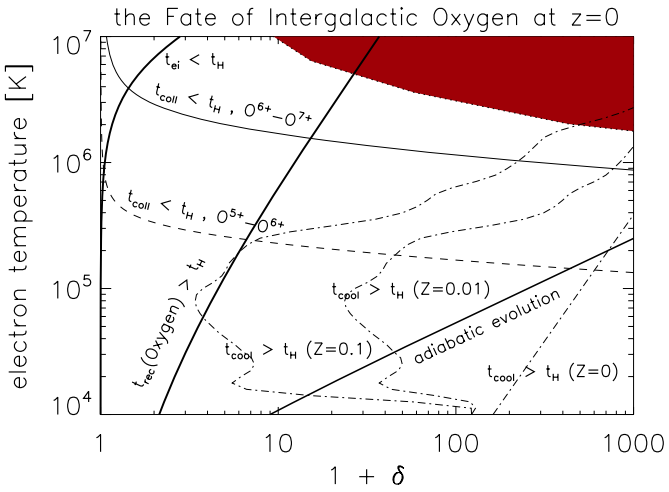


Fig. 4 Phase diagram for oxygen in the IGM at redshift $z = 0$. Density is parameterised as $\delta = \rho/\langle\rho\rangle$. The various boundaries separate regimes under which a certain process does, or does not attain equilibrium over a Hubble time. Critical boundaries for kinetic and thermal equilibrium are: the *solid line* in the upper left hand corner labelled '$t_{\rm ei} < t_{\rm H}$' indicates where electron and proton fluids reach kinetic equilibrium (proton temperature equal to electron temperature) in a Hubble time ($t_{\rm H}$): at low density and high temperature, such equilibrium does not obtain. Low density gas will not radiatively cool over a Hubble time to the *left* of the boundaries marked '$t_{\rm cool} > t_{\rm H}$'; the three curves are labelled with the metallicity, Z, expressed as a fraction of Solar metallicity. The cooling time was calculated for collisionally ionised gas. The *solid curve* labelled '*adiabatic evolution*' indicates the locus of gas that has only undergone adiabatic compression or expansion since high redshift (initial condition $T_{\rm e} \sim 10^4$ K); all shock-heated gas will be above this line right after passing through a shock. Critical boundaries for the ionisation equilibrium of oxygen are: the *shaded area* in the upper right hand corner indicates the regime where the collisional ionisation timescale is shorter than the photoionisation timescale, for ionisation O VIII→O IX ($O^{7+} \rightarrow O^{8+}$). The two boundaries labelled '$t_{\rm coll} < t_{\rm H}$' indicate where the collisional ionisation timescale becomes shorter than the Hubble time. At lower temperature, the ionisation balance cannot be in (collisional) equilibrium. *Upper* (*solid*) *curve* is for ionisation O VII→O VIII ($O^{6+} \rightarrow O^{7+}$), *lower* (*dashed*) *curve* for O VI→O VII ($O^{5+} \rightarrow O^{6+}$). The steep solid curve labelled '$t_{\rm rec}$(Oxygen) > $t_{\rm H}$' indicates where the radiative recombination timescale (O IX→O VIII or $O^{8+} \rightarrow O^{7+}$) exceeds the Hubble time (no recombination at low densities)

2.4 Spectroscopy: A Warning

It has been emphasised throughout the papers in this issue that advancing our understanding of the low-density IGM (outskirts of clusters and the WHIM) means: determining the physical properties (density, temperature, abundances, possibly velocity fields) of the medium to an accuracy sufficient to uniquely trace total mass and thermodynamical/enrichment history. A mere detection, and a rough determination of the total mass density in the WHIM, is not sufficient motivation to warrant the (potentially large) expenditure of funds and other resources on the experiments required to address this problem. Also, a proper study of the astrophysics of the low-density, highly ionised IGM naturally complements studies in other fields (galaxy formation and evolution, signatures of cosmic feedback from star formation and supermassive Black Holes, maybe even reionisation) and therefore forms a natural component of modern astrophysics.

We therefore emphasise again the importance of high resolution spectroscopy, spatial resolution, and sky coverage. Below, we reproduce the phase diagram of metals in the low-density IGM at $z \sim 0$, as shown in Bykov et al. (2008), to emphasise the fact that mea-

surements and conclusions based on simple assumptions concerning the physical state of the metals in the highly ionised IGM are likely to yield biased and unreliable results. For instance, even detection of both an O VII $n = 1$–2 resonance absorption line as well as a corresponding O VIII Lyα absorption line from a single filament, does not uniquely translate into an electron temperature, unless one knows what the dominant ionisation mechanism is. And in fact, it is quite possible that the ionisation balance between H- and He-like oxygen has not reached equilibrium in the filament under study, if its density is very low. Likewise, at limited spatial resolution, we may confuse the topology of line emitting and absorbing gas, relative to the distribution of galaxies, and our images are confused by metals freshly injected into the IGM by actively star-forming galaxies. Finally, at limited sensitivity, we will collect a sample of absorption line measurements and emission line images that will necessarily be biased towards the most massive filaments (the emission line images even more so than the absorption line detections), as well as the most nearby ones (in emission). Given the very large spread in density, temperature, and enrichment expected in the WHIM, we should keep in mind that sampling the properties of the most massive, densest phases of the medium is highly biased, and our measurements should reach well below these 'tips of the iceberg' before we can claim to have characterised the WHIM.

3 Hard X-ray and γ-ray Emission

We conclude with a brief, non-exhaustive description of future instrumentation for the detection of hard X-ray and γ-ray emission associated with nonthermal processes, as described in Petrosian et al. (2008)—Chap. 10, this issue. The potential of *GLAST* for detection of high energy γ-rays is briefly addressed in this latter reference; here, we concentrate on detection of hard X-ray emission (traditionally approximately the 10–100 keV band). The challenge is to detect faint, extended diffuse emission over a large ($\lesssim$ several arcminutes) field of view in the presence of strong detector backgrounds. Traditionally, instruments in this band have operated with collimators or coded masks, because it is difficult to focus hard X-rays for true imaging. However, the coming generation of hard X-ray experiments is built around innovative focusing optics, which promises much higher contrasts for astrophysical sources, and the detection of hard nonthermal emission from clusters is one of the primary science areas for these new instruments.

Simbol-X is a hard X-ray mission, operating in the 0.5–80 keV range, proposed as a collaboration between the French and Italian space agencies with participation of German laboratories for a launch in 2013 (Ferrando et al. 2006). Relying on two spacecraft in a formation flying configuration, *Simbol-X* uses a 20–30 m focal length X-ray mirror to focus X-rays with energies above 10 keV, resulting in over two orders of magnitude improvement in angular resolution and sensitivity in the hard X-ray range with respect to non-focusing techniques. The field of view of the instrument (at 30 keV) is $> 12'$ in diameter. The on-axis sensitivity is better than 10^{-14} erg cm^{-2} s^{-1} (i.e. 0.5 μCrab) in the 10–40 keV band, providing a 3σ detection in 1 Ms for a power law source spectrum with a photon index of 1.6. The goals for on-axis effective area S are: at 0.5 keV photon energy S > 100 cm^2, at 2 keV S > 1000 cm^2, at 8 keV S > 600 cm^2, at 30 keV S > 300 cm^2, at 70 keV S > 100 cm^2, and at 80 keV S > 50 cm^2.

The focal plane has two detector systems: the Low Energy Detector (LED) and the High Energy Detector (HED). The heart of the low energy detector is a monolithic macro pixel detector array of 128 $\times$ 128 pixels. The readout scheme of the pixel detector is that of an active pixel sensor. An active pixel sensor consists of a Silicon drift detector with an

integrated DEPFET for readout. An optical blocking filter is deposited directly onto the surface of the LED. The LED detector is made of a single Si wafer on which the 128 × 128 pixels are integrated. The HED is a hard X-ray camera made of an array of modules of 1 cm^2 arranged in a 8 × 8 square to cover the field of view. The detector material is CdTe (or CdZnTe) which provides the stopping power necessary to detect the hard X-rays. This material has flight heritage on *INTEGRAL* and *SWIFT*. The broad sensitivity calculated for the 10–40 keV band provides a flux limit of about 6×10^{-15} erg cm^{-2} s^{-1} that can be reached in 1 Ms.

The *NuSTAR* (Harrison et al. 2005) and *NEXT* missions (Nakazawa et al. 2006) similarly feature focusing hard X-ray optics.

Acknowledgements The authors express their gratitude to ISSI for its support of the team 'Non-virialized X-ray components in clusters of galaxies'. FP acknowledges support from the Dutch Organisation for Scientific Research NWO, and expresses his gratitude to SRON, the Netherlands Institute for Space Research, for its hospitality.

References

X. Barcons, XEUS—Studying the evolution of the hot Universe. MPE Rep. **281**, 77 (2003)
A.M. Bykov, F.B.S. Paerels, V. Petrosian, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9309-4
E. Churazov, M. Haehnelt, O. Kotov, R. Sunyaev, Mon. Not. R. Astron. Soc. **323**, 93 (2001)
J. Cottam, W. Cash, K. Flanagan et al., Proc. SPIE **6266**, 62661X (2006)
R. Davé, R. Cen, J. Ostriker et al., Astrophys. J. **552**, 473 (2001)
J.-W. den Herder, J.S. Kaastra, F.B.S. Paerels et al., Proc. SPIE **6266**, 62660H (2006)
M. Elvis, F. Fiore, Proc. SPIE **4851**, 377 (2003) (astro-ph/0303444)
C. Enss, *Cryogenic Particle Detection* (Springer, Heidelberg, 2005)
P. Ferrando, M. Arnaud, U. Briel et al., Proc. SPIE **6266**, 62660F (2006)
M. Fukugita, P.J.E. Peebles, Astrophys. J. **616**, 643 (2004)
F.A. Harrison, F. Christensen, W. Craig et al., Exp. Astron. **20**, 131 (2005)
U. Hellsten, N.Y. Gnedin, J. Miralda-Escudé, Astrophys. J. **509**, 56 (1998)
J.S. Kaastra, F.B.S. Paerels, F. Durret, S. Schindler, P. Richter, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9310-y
D. McCammon, W. Sanders, Annu. Rev. Astron. Astrophys. **28**, 657 (1990)
D. McCammon, R. Almy, E. Apodaca et al., Astrophys. J. **576**, 188 (2002)
R. McEntaffer, S. Osterman, W. Cash et al., Proc. SPIE **5168**, 492 (2004)
R. Mewe, in *X-ray Spectroscopy in Astrophysics*, ed. by J. van Paradijs, J.A.M. Bleeker (Springer, Heidelberg, 1999)
K. Nakazawa, Y. Fukazawa, T. Kamae et al., Proc. SPIE **6266**, 62662H (2006)
T. Ohashi, M. Ishida, S. Sasaki et al., Proc. SPIE **6266**, 62660G (2006)
V. Petrosian, A.M. Bykov, Y. Rephaeli, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9327-2
L. Piro, L. Amati, M. Barbera et al., Proc. SPIE **6266**, 62660K (2006)
L. Piro, J.-W. den Herder, T. Ohashi et al., 2007, ESA Cosmic Vision proposal (astro-ph/07074103)
A.P. Rasmussen, S.M. Kahn, F. Paerels et al., Astrophys. J. **656**, 129 (2007)
P. Richter, F.B.S. Paerels, J.S. Kaastra, Space Sci. Rev. (2008). doi:10.1007/s11214-008-9325-4
B.J. Wargelin, M. Markevitch, M. Juda et al., Astrophys. J. **607**, 596 (2004)
R.J. Williams, S. Mathur, F. Nicastro et al., Astrophys. J. **631**, 856 (2005)